LIBRARY

DEC 29 1925

Vol. VIII, No. 6. *PSYCHOLOGICAL REVIEW PUBLICATIONS* DECEMBER, 1925

Journal of Experimental Psychology

EDITED BY

JOHN B. WATSON, NEW YORK CITY
HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*Index*) AND
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)

ADVISORY EDITORS

H. A. CARR, UNIVERSITY OF CHICAGO; RAYMOND DODGE, WESLEYAN UNIVERSITY; KNIGHT DUNLAP, THE JOHNS HOPKINS UNIVERSITY; H. S. LANGFELD, PRINCETON UNIVERSITY; E. L. THORNDIKE, COLUMBIA UNIVERSITY.

CONTENTS

Judgments of Certain Space Relations Based upon the Learning of a Stylus Maze: CARL JOHN WARDEN 399

An Experimental Study of the Seashore Consonance Test: CHRISTIAN PAUL HEINLEIN 408

The Spatial Threshold of Touch in Blind and in Seeing Children: MARGARET S. BROWN AND GEORGE M. STRATTON 434

The Development of Number Concept in Children of Pre-School and Kindergarten Ages: HARL R. DOUGLASS 443

Instrument for Marking the Temperature Spots on the Skin: C. R. PENDLETON 471

PUBLISHED BI-MONTHLY BY
PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STREETS, LANCASTER, PA., AND
PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

Psychological Review Publications

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
JOHN B. WATSON, 244 MADISON AVE., NEW YORK (*J. of Exp. Psych.*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*Index*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)
WITH THE CO-OPERATION OF
MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL BULLETIN

containing abstracts, critical reviews, of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in May, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

Philosophical Monographs: a series of treatises more philosophical in character.
Library of Genetic Science and Philosophy: a series of bound volumes.

ANNUAL SUBSCRIPTION RATES

Review: $5.00 (Foreign, $5.25). **Review and Bulletin:** $9.50 (Foreign, $10.00)
Journal: $5.00 (Foreign, $5.25). **Review and Journal:** $9.00 (Foreign, $9.50)
Bulletin: $5.50 (Foreign, $5.75). **Journal and Bulletin:** $9.50 (Foreign, $10.00)
Any one of above with **Index:** $1.50 additional.
Review, Bulletin, and Journal: $14.00 (Foreign, $14.75).
Review, Bulletin, Journal, and Index: $15.00 (Foreign, $15.75).
Current Numbers: Review or Journal, $1.00; Bulletin, 60c; **Index,** $2.00
Psychological Monographs: $6.00 per volume (Foreign, $6.30).
Current Issues: prices vary according to size.

Subscriptions, orders, and business communications may be sent direct to the

PRINCETON, N. J. OFFICE

PSYCHOLOGICAL REVIEW COMPANY

FOREIGN AGENTS: G. E. STECHERT & CO., London (2 Star Yard, Cary St., W.C.) Paris (16, rue de Condé)

ExP

Journal of
Experimental Psychology

VOL. VIII., No. 6 December, 1925

JUDGMENTS OF CERTAIN SPACE RELATIONS BASED UPON THE LEARNING OF A STYLUS MAZE

BY CARL JOHN WARDEN

Columbia University

This study is a preliminary attempt to determine to what extent the non-visual experience involved in the learning of a stylus maze can be utilized, immediately afterward, as the basis of judgments concerning the spatial character of the maze pattern, in terms of our usual system of visual measurements.

The data were gathered in connection with the mastery of a maze of the simple design of Fig. 1 by 60 subjects, most of whom were students of Columbia College, and none of whom had had previous experience in the stylus maze situation. The maze was learned under standard conditions,[1] and the subject was given no suggestion whatsoever in the instructions that he would be required, after having mastered the maze, to give information concerning the spatial character of the pattern.

The maze measured $10\frac{1}{4}$ x $11\frac{1}{4}$ inches, the longer direction extending forward from the subject when seated for practice. All visual cues were eliminated by screening the entire table on which the maze rested by means of a hood of black cloth which extended about 20 inches above the table. The size of the hood and table top (24 x 36 in.) gave the subject some

[1] For details of method employed see article by the writer, JOUR. EXPER. PSY., VII., August 1924, 243–275.

indication of the size of the maze itself, since these dimensions would obviously mark the upper limits of its possible dimensions. Judgments concerning the size of the maze were evidently influenced, in a few cases at least, by this factor.

After the maze had been learned, the following procedure was carried out in the order detailed below: (1) The subject was required under non-visual conditions, to reproduce the maze pattern (true pathway) on a large sheet of graph paper placed under the hood, before shifting the bodily posture

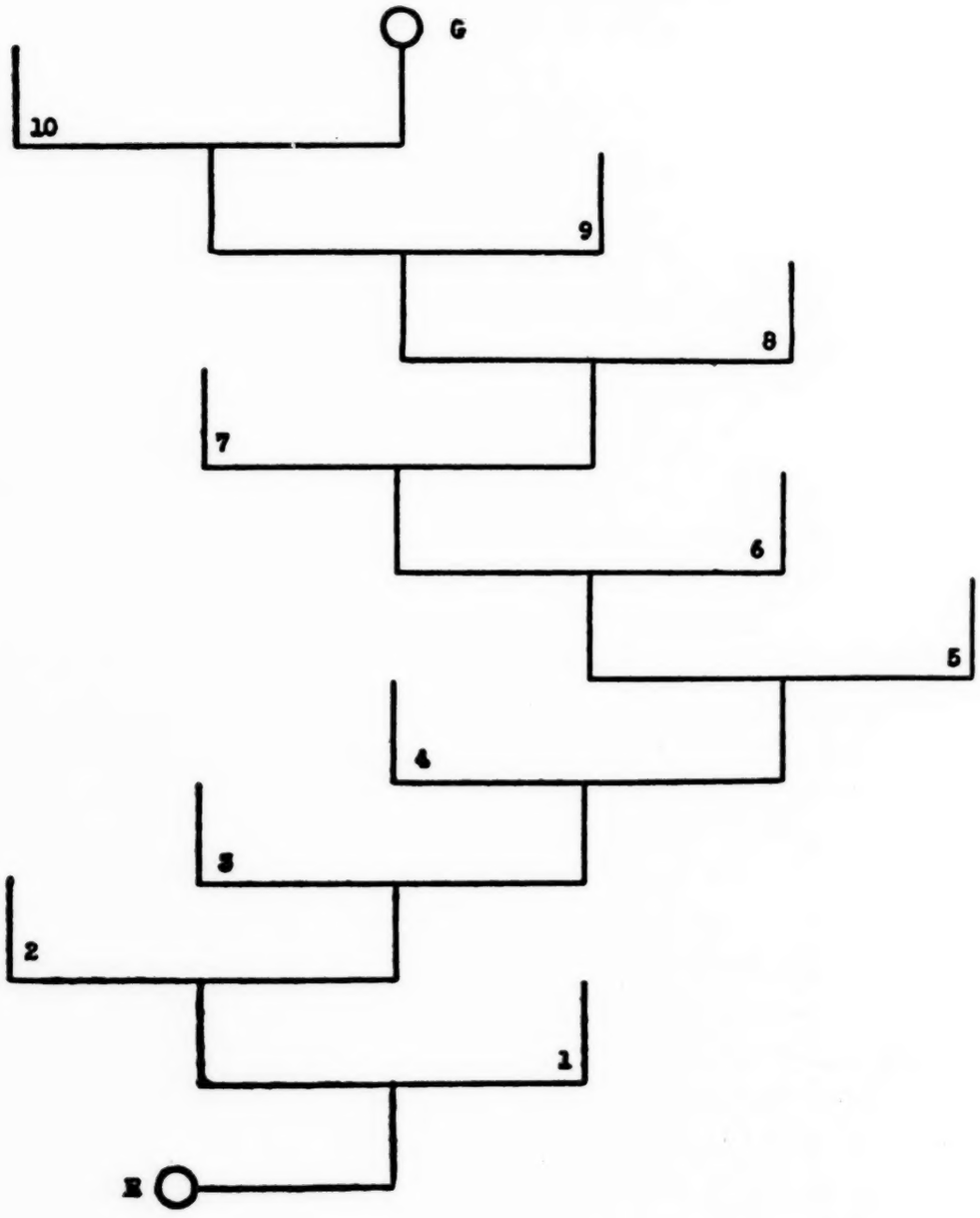


FIG. 1

from that assumed during practice. (2) On a similar sheet the subject was instructed to draw the entrance and goal areas, in the order named. The two areas were exactly the same size being circular in form with a diameter of ¾ of an inch. (3) The subject then turned to a table and drew the same two areas, this time with vision. (4) A further drawing was now required indicating the shape of the cul de sac most frequently encountered in the maze. (5) A series of judgments or estimates concerning the spatial character of the pattern were then entered by the subject on mimeographed sheets.

The results obtained in the manual reproduction of the maze pattern have been reported in a previous paper (supra, p. 254 ff.). The remainder of the data will be presented topically below, following the chronological order in which it was gathered.

Results

I. *Manual vs. Manual-Visual Reproduction of Entrance and Goal Areas.*—The specific instruction to the subject, in the manual reproduction, was as follows: "Move your hand, containing the pencil, as if you were moving the stylus around the rim of the entrance (or goal) area—I want to see how large it seems to you." The factor of size-estimation was thus emphasized in the manual as well as the manual-visual reproduction.

A distribution of the circles drawn to represent the entrance and goal areas, in terms of their diameters, is shown in Table I. The diameter recorded for each drawing was obtained by averaging the longest and shortest distances across the more or less circular outline, measured to the sixteenth of an inch.

Table I

Distribution for Size of Reproduced Entrance and Goal Areas

	¼– ½ in.	½– ¾ in.	¾– 1 in.	1– 1½ in.	1½– 2 in.	2– 4 in.	Ave. Dia.
Manual (*E*)	3	16	24	14	3	0	15/16 in.
Man-vis. (*E*)	6	7	18	19	6	4	18/16 "
Manual (*G*)	10	14	19	15	1	1	14/16 "
Man-vis. (*G*)	11	8	10	21	5	5	17/16 "

The actual diameter of both *E* and *G* circular areas was ¾ of an inch. The 60 subjects are distributed in the table on the basis of the average diameter of the area reproduced under the different conditions.

The range in size of the reproduced areas is extremely large as indicated by the headings of the table. In several instances the diameters were drawn as small as ¼ inch, while the other extreme, in which the diameter measured 3 3/5 inches, or nearly five times that of the actual area, occurred twice.

The averages of the last column indicate a general tendency to overestimate in both types of reproduction. The same fact appears in the distribution for size. Only 17 of the total of 240 drawings were exact reproductions, and only 58 involved underestimation. The remaining 165 cases were overestimations. Of these latter, the diameter was drawn at least twice as long as that of the actual area, in over 15 per cent. of the cases. The goal area was estimated to be somewhat smaller than the entrance, the very small size occurring the more frequently. If the difference is at all significant it might possibly be accounted for by the fact that, in learning the maze, the arm is more extended in traversing the goal area. There is likely also to be more exploration in the entrance area which should effectually offset extreme underestimation.

Overestimation occurred more frequently and to a greater degree in the manual-visual than in the manual reproduction. This was true for both entrance and goal areas. The average diameter for the group was approximately 20 per cent. greater in the manual-visual series for both areas. The modal diameter of the manual series lies between the limits ¾ to 1 inch for both *E* and *G*, while that of the manual-visual series lies between 1 to 1½ inches. The manual-visual series shows also more cases of extreme deviation from the mode, especially in the direction of the larger sizes. The data indicate very definitely that non-visual reproduction is more accurate than visual, when the only contact with the spatial situation has been also non-visual, and more or less incidental as in the present case. The reliability of the difference between the two modes of reproduction is very high, the difference being more than 20 times the S.D. of the difference in both the *E* and *G* series.

The above comparisons are based upon the group results and hence show nothing as to the consistency of the individual in making the four required reactions. In general, the error of estimation lay in the same direction, although the amount often differed greatly, in the entrance and goal drawings under the two conditions. That is, if a subject drew the entrance larger without vision he would likely draw the entrance larger also when vision was used. In only 25 per cent. of the cases was this relationship reversed, and in half of these instances the *E* and *G* drawings were made merely equal in either the visual or non-visual series. There was considerable overlapping between the two manual and the two manual-visual series, because of the very large difference between the entrance and goal drawings of either series. That is, the difference between *E* and *G* was usually greater than that between the two *E*'s (manual and manual-visual) or the two *G*'s. For example, a record such as the following was not unusual: Manual-visual series (*E*, 24/16 in., *G*, 12/16 in.) and Manual series (*E*, 14/16 in., *G*, 9/16 in.). The absolute differences for a given individual were often very large.

II. *The Form of the Culs-de-sac.*—An attempt was made to find out what notion the subject had formed of the shape of the blind alleys during the learning process. He was required to draw, with the aid of vision, the form of alley which 'he remembered best,' ignoring the size aspect altogether.

The maze included only a single type, shaped like a capital L, the approach being through the longer segment (4 in.) and into the shorter segment (2 in.) by a right angle turn. Of the sixty subjects, only 20 drew an L-shaped cul-de-sac, the remaining 40 drawing a simple straight line, evidently corresponding to the shorter segment of the *L*, except in the case of one subject who drew a short segment of a circle, explaining that "all the grooves in the maze were arcs of a circle" similar to the one he had drawn. It is evident that the majority of the subjects had not discovered that the cul-de-sac was L-shaped, even though the blind alleys had been either completely or partially traversed on an average of about 70 times during the period of mastery, and in certain

cases as many as 400 to 500 times. The longer segment of the cul-de-sac seemed to be neglected in the drawing. Evidently few of the subjects had discovered that there was only a single type of alley in the maze, since they usually asked if they should draw "some of the other forms" also, in addition to the one best remembered.

III. *The Number and Position of Culs-de-sac in the Maze.*—After learning the maze the subjects were required to make an estimate of the number of blind alleys to the right and to the left of the true pathway. In most cases the subject reported a larger number on one side than on the other, only 11 estimating an equal number on both sides. The group average was 3.2 culs-de-sac for the right side and 3.5 for the left side, the difference being obviously insignificant.

The data concerning the total number of alleys in the maze are more interesting. Only 2 of the 60 subjects reported the correct number (10) whereas 4 overestimated the number. Ninety per cent. of the group underestimated the number and 35 per cent. reported only half the correct number, or less. It should be noted that, of the 60 subjects, 43 had actually entered all of the culs-de-sac at some time during the period of learning, 15 had entered all but one, while the remaining 2 subjects had entered all but two of the ten.

TABLE II

DISTRIBUTION FOR ESTIMATION OF NUMBER OF CULS-DE-SAC IN MAZE

Number Estimated.....	3	4	5	6	7	8	9	10	11	12	14	Av. 6.5
Number of subjects.....	1	2	14	11	9	8	5	2	2	1	1	60 subj.

No doubt the general tendency to underestimation was due in part to the fact that all 10 culs-de-sac were seldom entered at a single trial. Certain alleys also seemed to have a greater position value than others and became the locus of dominant errors. Confusion of closely contiguous alleys with these more definitely localized ones was evident. For example, number 3 was often confused with number 4 and number 9 with the goal area. Such confusions would tend to reduce

the estimated number, since a newly entered alley would seem to be identical with a contiguous one encountered on some previous trial.

IV. *The Relative and Absolute Size of the Maze.*—The subject was now required to make a judgment as to whether the forward or the lateral distance across the maze was the greater. The actual maze measured 10¼ inches lateral by 11¼ forward, the latter distance being exactly 1 inch greater than the former. Fifteen of the 60 subjects judged the lateral distance to be greater, the error in absolute terms ranging from 1½ to 9 inches with an average of 6.3 inches. Eleven subjects judged both distances to be equal. The remaining 34 subjects correctly judged the forward distance to be greater than the lateral, but greatly overestimated the amount of the difference, the range of error being from ½ inch to 12 inches, with an average of approximately 4 inches.

After the experiment was well under way, it became fairly obvious from the extremely large errors occurring in the judgments just noted, that the maze must seem very much larger to the subjects than it actually was. For, in certain of the above judgments, the difference between the length and width was judged to be greater than either dimension, and 10 or 12 times the actual difference. The 41 remaining subjects were now required, in addition to the above exercise, to estimate in terms of inches, or fractional parts of an inch, the two dimensions of the maze.

The range was remarkably wide, the smallest size reported being 6 x 8 inches and the largest 24 x 36 inches, both limits being given by two different subjects. Estimates larger than 24 x 36 could hardly have been obtained, under the experimental conditions, since the hooded table on which the maze rested was also this size. In both cases in which the 24 x 36 estimate was given, the subject volunteered the information that it had seemed to him as if the maze covered the entire top of the table when learning it. This error, even though so large as to appear quite ridiculous, was not much greater than that made by 20 per cent. of the subjects, and is in agreement with the general tendency to overestimate the size of the maze.

In 18 out of the 41 cases the length of the maze was given as 18 inches or over, and in 16 instances the width was estimated at 15 inches or over. The average size of the maze for the group was 14.4 x 16.2 inches. In only 6 cases were both dimensions of the maze underestimated, and the amount of underestimation was comparatively small in every case. In 3 instances of underestimation of the lateral dimension, some one of the 3 outlying culs-de-sac had not been traversed by the subject during the period of mastery, and hence the full lateral distance had never been experienced kinesthetically. This fact may account for the error for the lateral measurement.

These results agree with those reported in section 1 above. There it was found that reproduction with vision of the entrance and goal areas, from non-visual experience only, resulted in at least a 20 per cent. overestimation. Here the large amount of non-visual contact in learning the maze (range of trials for group 16–195) resulted in even greater overestimation in translating the non-visual experience into units of a common visual measure. These facts would seem to emphasize the importance of the usual visual check on kinesthetic data utilized in space perception, and more especially in motor adjustment.

V. *Position of Goal Relative to Entrance.*—The problem at this point was to judge whether the goal was located to the right or to the left of the entrance, and to estimate in terms of inches the amount of the distance in either direction. The goal was actually located 2 inches to the right of the entrance (pattern 1) or 2 inches to the left when the pattern was reversed by turning the maze over. Thirty subjects learned the maze in each position.

Only about 60 per cent. of the combined group (60 subjects) were correct in their judgment as to the relative position of entrance and goal. About 15 per cent. placed the goal in a direct line forward from the entrance. The remaining 25 per cent. made an error of more than 2 inches and in the wrong direction.

The position assigned to the goal by the group average was

not far removed from the true position. The 30 subjects who learned pattern I., placed the goal 1.73 inches to the right of the entrance, the true position being 2 inches to the right. The same value for the group using pattern II. was 0.84 inches to the left, the true position being, in this case, 2 inches to the left.

The apparent accuracy of judgment indicated in the preceding paragraph is highly fictitious. For the group learning pattern I., the goal position for different individuals fell between the limits 3 inches left to 8 inches right. For the other group the range of position was from 6 inches to the right to 6 inches to the left. In both cases, then, the range of position assigned to the goal by members of the groups, was greater than the lateral dimension of the maze. In fact 13 per cent. of the subjects located the goal at a point that lay entirely outside the maze pattern. The extreme inaccuracies observed in the making of this simple judgment of relative position seem to be due very largely to overestimation of movements made under non-visual conditions.

In conclusion it might be remarked that when the subjects were shown the maze, after having learned it and given the information here reported concerning its spatial character, they usually expressed surprise. Either the size, or the arrangement of culs-de-sac, or the relative position of entrance and goal areas, or some other feature of the pattern differed from that which the subject had expected. It would be interesting to know whether or not more accurate estimates could be made in a similar situation if the subject were instructed beforehand to attempt to translate his movements, during the learning process, into our usual system of visual measurements.

AN EXPERIMENTAL STUDY OF THE SEASHORE CONSONANCE TEST

BY CHRISTIAN PAUL HEINLEIN

From the Psychology Laboratory of the Johns Hopkins University

I

It is a generally recognized fact among musicians that dissonance is closely associated with a feeling of unrest or tension, whereas consonance is associated with a feeling of rest or repose. Indeed, so prominent a rôle does this tension-repose dimension of feeling play in consonant evaluation, that it is extremely difficult to deny its existence in a test such as that formulated by Carl E. Seashore and his collaborators (1; 3; 6). Dissonance, since it arouses feelings of dissatisfaction, implies necessary motion towards something of a more satisfying nature: *i.e.*, motion towards consonance. This passage from dissonance to relative consonance may be referred to as *resolution.* Such a theory which bases the necessity for resolution on aesthetic principles is by no means a new one. Helmholtz (2, p. 331) proposes just such a notion when he points out that dissonances are used partly as a "means of contrast, to give prominence to the impression made by consonances, and partly means of expression, not merely for peculiar and isolated emotional disturbances, but generally to heighten the impression of musical progress and impetuosity, because when the ear has been distressed by dissonances it longs to return to the calm current of consonances."

Pole (5, p. 272) champions the same view when he states that "there is every reason to believe that the feeling of the necessity for the resolution of a discord arises simply from the natural desire of the mind to *rest upon agreeable impressions*, rather than on disagreeable ones." In seeking further support for this view, Pole quotes Moritz Hauptmann who declares

that "a dissonance requires a resolution for the non-musical as well as for the musical hearer; a dissonance is something unmeaning for every ear."

It is here quite needless to enter into any prolix discussion of the psychological significance of harmonic relationship. Much has been written on the topic, yet little of the written material is conclusive. The principal point which here needs immediate emphasis is the fact that in Seashore's consonance test definite harmonic relationships exist, and that the nature of the accompanying feeling-tone is largely determined by these relationships. Indeed, it seems that Seashore has largely disregarded the various harmonic relationships which have entered into his test-form. Realization of the danger of musical progression, however, is indicated by Malmberg, one of Seashore's co-workers, who says:

> Restfulness—a feeling of completeness, finality, or satisfaction, with its opposite disquietude—a feeling of incompleteness, needing to be resolved, was first adopted as a fifth criterion, but it soon developed that it must be dropped as it is a variable criterion directly due to progression and association, which must be excluded (3, p. 108).

The grievous discrepancy here lies in the fact that a simple assertion henceforth dispenses with such undesirable factors with which the adopted criteria may conflict. Thus the exclusion takes upon itself the nature of a mere verbal statement: it is not by any means an actuality.

It should be recalled that a rest of but one second separates the musical intervals. In fact, such a brief pause between intervals makes resolution the more obvious. Nor is the brief pause of approximately two to four seconds between pairs of sufficient duration to destroy any awareness of harmonic progression which may exist between the last and first binary combinations of two consecutive pairs. That harmonic progression plays a dominant rôle in the conscious reaction of the musician to any series of intervals, paired or otherwise, is not to be doubted. Through training and continued practice, the musician has developed and cultivated an entire system of auditory habits and associations which is not readily subject to modification or repression at any single sitting, in spite of all attempts to coöperate on the

part of the subject. Rigidity of instructions in the form of nicety of statement is insufficient to eliminate to any appreciable extent the feeling elements which are so firmly linked with these auditory habits and associations. Under such conditions, one is not surprised in finding such marked absence of correlation between musical training or ability to perform and perception of consonance based on the proposed criteria and test material (3, p. 127). Indeed, the test-form seems to engender those very factors of normal auditory experience—feeling-tone, association, and musical relationship —which Seashore has endeavored to dismiss.

Peterson (4, p. 31) in speaking of Seashore's consonance test says:

> Tonic effects, preparations for resolutions, associations with this or that progression in familiar musical selections, etc., are almost inevitable and seem to be troublesome to many subjects. There is hardly such a thing as a purely independent comparison of two sets of intervals on their own degrees of consonance. This aspect of the subject needs careful investigation.

II.

The writer has made an experimental investigation for the purpose of ascertaining the extent of influence which harmonic principles and the various laws of musical progression may have upon judgment in the paired interval comparison method employed by Seashore in his consonance test.

On February 16, 1925, at 11:30 A.M., the Seashore consonance test was given to thirty-five male college students representing a class in General Introductory Psychology in the College of Arts and Sciences at the Johns Hopkins University. Each student was supplied with a mimeographed test blank upon which he recorded his judgments. An electrically-driven console model Columbia phonograph was employed for the test. The instrument was set for seventy-eight revolutions per minute. "Loud" steel needles were used. Each student was required to fill out a questionnaire relating to his past and present musical experience and training. This questionnaire was a modification of the one used by Seashore (7). Through careful scrutiny of the responses given, it was quite possible to ascertain roughly

certain facts relating to each subject's musical training, musical environment, musical expression, and musical appreciation.

Before the disc was played, the instructions (8, p. 14) were carefully and distinctly read to the listeners. The instructions were repeated in order to make certain that they were understood. No general discussion nor further explanation followed the reading of the instructions. No preliminary practice was afforded.

At the end of the test, the students were allowed a rest of approximately ten minutes, during which time many completed the filling out of their questionnaires which had not been completed before the test began. At all events, the students were directed to refrain from conversation with their neighbors.

After the ten-minute intermission, the instructions were again read and the test repeated. Through a conspicuous shifting of records in the album, the director purposed to lead the subjects to believe that the repetition was actually a different test-record. Moreover, the blanks used by the students during the first test were marked in an 'A–B' form: the blanks used during the second test were marked in a 'C–D' form. After the record began, to what degree or how long this illusion was entertained on the part of the subjects, if it was entertained at all, could not be determined. This illusion was purposely designed to prevent the subjects from consciously attempting to recall any of their previous judgments which they had made during any portion of the first test. The success of technique involving purposeful dissimulation, despite its nicety of application, can only be conjectural.

The results of the tests were analyzed and studied for the following points:

(1) The total number of errors for each subject, and the particular intervals which were misjudged.
(2) The occurrence in judgment of different types of reversals for any particular interval.

(3) The relation of intervals within a single pair in which reversal of judgment had occurred.
(4) Consistent erring in the same interval upon its re-occurrence in the same test.
(5) A comparison between the first test and its repetition in the number and types of errors made.

The data compiled and analyzed were based on a system of preferences. When the subject recorded that the second interval was 'worse' than the first interval, he thereby implied that he 'preferred' the first interval. Preference for the second interval was directly indicated by a judgment of 'better.'

TABLE I

INTERVAL VALUES REPRESENTED IN SEASHORE'S CONSONANCE TEST

Pairs are grouped to illustrate interval reversals. An asterisk before an interval represents, according to Seashore, the greater consonance or less dissonance of two intervals.

Pair	Intervals		Pair	Intervals	
1....	* f′a′	Maj. 3d	7...	f′a′	Maj. 3d
	b♭a′	Maj. 7th		* c′a′	Maj. 6th
40....	b♭a′	Maj. 7th	34...	* c′a′	Maj. 6th
	* f′a′	Maj. 3d		f′a′	Maj. 3d
2....	* a♭d′	Aug. 4th	8...	* g b	Maj. 3d
	a♭g′	Maj. 7th		g e′♭	Min. 6th
39....	a♭g′	Maj. 7th	33...	g e′♭	Min. 6th
	* a♭d′	Aug. 4th		* g b	Maj. 3d
3....	e′♭b′♭	Perf. 5th	9...	* c′b′♭	Min. 7th
	* b♭b′♭	Octave		c′b′	Maj. 7th
38....	* b♭b′♭	Octave	32...	c′b′	Maj. 7th
	e′♭b′♭	Perf. 5th		* c′b′♭	Min. 7th
4....	g♯b	Min. 3d	10...	f′♯g′	Min. 2d
	* g♯e′	Min. 6th		* f′g′	Maj. 2d
37....	* g♯e′	Min. 6th	31...	* f′g′	Maj. 2d
	g♯b	Min. 3d		f′♯g′	Min. 2d
5....	d′a′♭	Dim. 5th	11...	* a c′	Min. 3d
	* c′a′♭	Min. 6th		b♭c′	Maj. 2d
36....	* c′a′♭	Min. 6th	30...	b♭c′	Maj. 2d
	d′a′♭	Dim. 5th		* a c′	Min. 3d
6....	* a g′	Min. 7th	12...	* a♭f′	Maj. 6th
	a b	Maj. 2d		a♭g′	Maj. 7th
35....	a b	Maj. 2d	29...	a♭g′	Maj. 7th
	* a g′	Min. 7th		* a♭f′	Maj. 6th

TABLE I. (*Continued*)

Pair	Intervals		Pair	Intervals	
14....	* c′b′♭	Min. 7th	20...	* e′♭g′♭	Min. 3d
	c′d′♭	Min. 2d		c′g′♭	Dim. 5th
27....	c′d′♭	Min. 2d	46...	c′g′♭	Dim. 5th
	* c′b′♭	Min. 7th		* e′♭g′♭	Min. 3d
15....	* g c′	Perf. 4th	21...	a e′♭	Dim. 5th
	b c′	Min. 2d		* a f′	Min. 6th
26....	b c′	Min. 2d	45...	* a f′	Min. 6th
	* g c′	Perf. 4th		a e′♭	Dim. 5th
16....	e′b′♭	Dim. 5th	22...	* b g′♯	Maj. 6th
	* e′g′♯	Maj. 3d		b a′	Min. 7th
50....	* e′g′♯	Maj. 3d	44...	b a′	Min. 7th
	e′b′♭	Dim. 5th		* b g′♯	Maj. 6th
17....	g a	Maj. 2d	23...	c′f′	Perf. 4th
	* g e′♭	Min. 6th		* c′e′	Maj. 3d
49....	* g e′♭	Min. 6th	43...	* c′e′	Maj. 3d
	g a	Maj. 2d		c′f′	Perf. 4th
18....	d′c′	Min. 7th	24...	* g g′	Octave
	* d′g′	Perf. 4th		g a	Maj. 2d
48....	* d′g′	Perf. 4th	42...	g a	Maj. 2d
	d′c′	Min. 7th		* g g′	Octave
19....	* a g′♯	Maj. 7th	25...	d′b′♭	Min. 6th
	f′×g′♯	Min. 2d		* d′b′	Maj. 6th
47....	f′×g′♯	Min. 2d	41...	* d′b′	Maj. 6th
	* a g′♯	Maj. 7th		d′b′♭	Min. 6th

Since the intervals of twenty-four[1] of the pairs which occur in the test are reversed in time-order, particular interval judgments have been studied in relation to their reoccurrence sometime in the test. (See Table I., which illustrates in detail Seashore's method of reversing intervals.) Thus, when, according to Seashore, the first interval of a pair and the second interval of the same pair reversed are the correct intervals, the following additional types of judgments in relation to the two pairs are possible:

I. When the subject prefers the first interval in each pair, the preference in the second pair being incorrect.

[1] The test includes 50 pairs: however, two pairs, nos. 13 and 28, are not reversed. Pair no. 13 is *a b♭–a a′*, a minor second and an octave. It is to be especially noted here that pair no. 28 on Seashore's record does not conform with the notation for that pair in Seashore's text (6, p. 155). Seashore's notation is *a a′–a f′*, an octave and a minor sixth; for the same pair, the phonograph disc records *a a′–a c′♯*, an octave and a major third.

II. When the subject prefers the second interval in each pair, the preference in the first pair being incorrect.
III. When the subject prefers the second interval in the first pair and the first interval in the second pair, both preferences being incorrect.

When, according to Seashore, the second interval of a pair and the first interval of the same pair reversed are the correct intervals, the following additional types of judgments in relation to the two pairs may occur:

IV. When the subject prefers the first interval in each pair, the preference in the first pair being incorrect.
V. When the subject prefers the second interval in each pair, the preference in the second pair being incorrect.
VI. When the subject prefers the first interval in the first pair and the second interval in the second pair, both preferences being incorrect.

In order to make such a list of judgment-types complete, account is taken of the occurrence of correct judgments in both pairs. Thus the following additional cases may occur:

VII. When the subject prefers the second interval in the first pair, and the first interval in the second pair, both judgments being correct.
VIII. When the subject prefers the first interval in the first pair, and the second interval in the second pair, both judgments being correct.

An analysis of these various judgment relationships is not without reason nor without prognostic value. If resolution does have its effect upon judgment, then certain types of judgment reversals can be expected to occur. Let us pause a moment and consider a particular illustration.

In passing from a major third to a perfect fourth within the same tonality, *e.g.*, c′e′–c′f′ (see pair no. 43), the ear passes harmonically from a dominant position to a tonic position as the easiest and most fundamental progression to a point of rest. The dominant position may be referred to as a 'floating' position: it gives an impression of incompleteness

because it causes an anticipation of completeness. Through this provided anticipation, it causes feelings of tension and strain which tend to persist until the ear is satisfied by the finality of the tonic. Melodically, the ear passes from the seventh or leading tone, which invariably demands motion toward finality, to the octave or point of rest. Physically, the progression from the major third to the perfect fourth is a passage from less to greater consonance. Now either of these consonant intervals, when isolated from the other, may be considered as points of repose. It is only when the order of paired interval comparison is employed that we gain a feeling of restfulness and greater finality in the last impression.

Now reverse the order—that is, let us pass from the perfect fourth to the major third; *e.g.*, c′f′–c′e′ (see pair no. 23). We are now confronted with the subdominant to tonic relation. The subdominant position in relation to the tonic is accompanied by feeling-tone analogous to the feeling-tone which accompanies the dominant in its relation to the tonic. In passing from the fourth to the third, the *f* is heard as leading downward into the third of the tonic; the third, therefore, gives the impression of completeness, despite the fact that physically the fourth is a greater consonance.

According to Seashore, the subject is expected to prefer the major third in each of the above pairs. But in experimentation it has been found that such preference is not consistently carried out. The frequency of certain types of reversals in judgment can be explained adequately only on a basis of harmonic relationship.

Of the 35 subjects tested, in the first test 13 subjects preferred the major third in pair no. 23 and the perfect fourth in pair no. 43. In the repetition of the test, 17 subjects made the same type of reversal. This number represents approximately one half of the total number of subjects. (Probable causes for increase of reversal in the second test over the first test will be listed later.) However, but 6 subjects consistently made this particular type of reversal in both tests. (Causes for such variability will also be discussed subsequently.) If a classification of judgment-types

were ignored, the number of subjects who preferred the fourth when it followed the third would be increased to 18 for the first test and 23 for the second test, with 8 subjects scoring such preference in both tests.

In pair no. 3, in the first test, 17 subjects preferred the octave to the perfect fifth, while in pair no. 38, the same 17 subjects preferred the perfect fifth to the octave. In the second test, 18 subjects made the same type of reversal. Twelve subjects consistently made the same reversal in both tests. Again the preference for the perfect fifth can be explained on a basis of harmonic relationship. Here we have a passage from the dominant to tonic through the melodic skip in the lower tones. It should be pointed out at this time that type of reaction may largely depend upon the degree of attention given to either the upper or lower tones of any particular combination or grouping of combinations.

In pair no. 4, in the first test, 20 subjects preferred the minor sixth to the minor third, while in pair no. 37, the same 20 subjects preferred the minor third to the minor sixth. In the second test, 27 subjects made the same type of reversal. Nineteen subjects consistently made the same reversal in both tests. Preference for the minor sixth in pair no. 4 can readily be explained on the basis of the melodic skip in the upper voice of the third to the tonic of the sixth. Preference for the minor third in pair no. 37 cannot be explained on a basis of resolution. The smaller interval was probably judged to be the better because it is a richer combination. This again seems to point to the fundamentality of feeling-tone in reaction. A subject in another group reacting to the test on another occasion informed the director after the test was given that she 'liked the smaller intervals better because they did not seem stretched out so far.'

Then again, melodic direction adds another complex. In descending from a relatively high tone to a relatively low tone, the low tone is almost invariably preferred (5, p. 275). In ascending, the low tone again is usually preferred, since heightened pitch causes a physiological strain upon the end-organ of hearing and consequently leads to a feeling of unpleasantness.

In pair no. 28, we have a situation similar to that found in pair no. 37. This pair represents a harmonic progression of the second order—the equivalent of a chord repetition which is in tonality but does not progress through tonality to any point of rest. In the first test, 19 subjects preferred the major third to the octave, while in the second test, 24 subjects preferred the major third to the octave. Seventeen subjects preferred the major third in both tests. Here again we have a marked descent to a relatively rich combination. Helmholtz (2, p. 370) writes:

> The resemblance of an Octave to its root is so great and striking that the dullest ear perceives it; the Octave seems to be almost a pure repetition of the root, as it, in fact, merely repeats a part of the compound tone of its root, without adding anything new. Hence the æsthetical effect of an Octave is that of a perfectly simple, but little attractive interval. The most attractive of the intervals, melodically and harmonically, are clearly the Thirds and Sixths,—the intervals which lie at the very boundary of those that the ear can grasp.

Thus the data presented repeat a fact which both Malmberg (3, p. 126) and Gaw (1, p. 140) discovered in an analysis of their own data. This adds another link of evidence in favor of the fundamentality of feeling-tone in auditory sensation.

The very nature of the test leads the subject to anticipate, through auditory imagery, a second interval before the second stimulus within a pair is actually provided. Indeed, it is likely that such anticipation greatly modifies judgment. If, when a combination is sounded, anticipation of some interval to follow is sufficiently vivid in imagination, then the anticipation, depending upon its nature and its relation to the given stimuli, may overshadow either of the intervals sounded, and a confused judgment follows as a result of comparison between the anticipated second interval and the second stimulus, or between the first stimulus and the anticipated second interval. Immediate preference for the first stimulus is a likely result when an anticipated second interval is unsatisfied by and conflicts with the second stimulus.

For an illustration, let us turn to pair no. 5. Here we have a diminished fifth preceding a minor sixth. Ordinarily, the diminished fifth is a somewhat disagreeable and un-

satisfying combination. It demands motion towards greater consonance as a point of rest. Now the minor sixth is unquestionably a greater consonance and may be considered as a point of repose if heard as a part of the tonic triad. However, in pair no. 5, the passage of d′a′♭ to c′a′♭ does not represent a very satisfactory resolution to the ear. Normally, the ear would expect resolution to e′♭g′; or, if d′a′♭ were conceived as d′g′♯, we might have resolution to c′♯a′. Either of these resolutions may be considered as primary or fundamental, since they both indicate passage from dominant to tonic. Thus, if these more fundamental resolutions are anticipated by the subject, marked confusion of judgment may result. Such confusion or quasi-comparison may account for the 19 subjects who preferred the diminished fifth to the minor sixth. However, it must be remembered that probabilities are here involved: there is no direct, objective method available for detecting such anticipations on the part of the subject.

At all events, it is safe to say that anticipation of a stimulus is very likely to modify judgment of the stimulus when it actually occurs. When confusion of judgment follows such quasi-comparison as already indicated, different attitudes necessarily follow, the attitudes themselves depending in a large measure on the nature and the relation of the combinations sounded, and on the nature of the anticipated interval and its relation to either or both of the combinations sounded.

When the second interval in a pair is a discord, the ear will refuse to accept the discord [2] as a satisfactory point of rest and may supply a resolution, in which case the comparison between the two intervals may be complicated by the presence of this imaged resolution which follows the second interval. This may explain, for example, why 18 subjects in the first test preferred the major second to the minor third in pair no. 11. In pair no. 30, the same 18 subjects preferred the minor third. This latter preference is to be expected in view of the fact that b♭c′ resolves satisfactorily into a c′. Preference for an interval may be expected

[2] Musically talented people frequently prefer discords to concords because of the progression possibilities which discords afford.

when the greater agreeableness which the interval creates coincides with a satisfactory harmonic progression. This is again illustrated in pair no. 34 in which case a total of 17 subjects in the first test and 23 subjects in the second test preferred the major third to the major sixth: 15 subjects consistently preferred the major third in both tests.

Peterson (4, p. 31), in referring to his investigation concerning reliability coefficients, writes:

> In view of the various attitudes developing from different sequences of intervals, it is interesting to compare the successive scores made by identical subjects in the Seashore consonance test which reproduces the immediately objective conditions of the several intervals about as constantly as machinery can do. Even under objectively controlled conditions of this kind various incidental irrelevant stimuli received during the tests, and the attitudes carried over from previous moments, are bound to have their effects, thus altering the constancy of a subject's responses to any given pair of intervals.

In the case of the 35 subjects under consideration, after the rest period of approximately ten minutes, repetition of the test under the same general conditions gave results which strongly confirm Peterson's assertion. Of the 35 subjects tested, 32 subjects scored a different grade [3] on the first test than on the repetition. Of these 32 subjects, 24 subjects showed an increase of errors in the repetition, while 8 subjects showed a decrease of errors in the repetition. Only 3 of the 35 subjects made the same number of errors in each test: however, in *none* of the individual records was the error distribution the same for both tests. This marked inconstancy between successive reactions is more clearly indicated in Table III. Notice especially the marked variation in percentile rank [4] of subjects 4, 9, 20, 27, 31, and 35. All of the subjects scored comparatively high on the first test, but dropped considerably on the repetition. The low ranking of subjects 5, 6, 10, and 12, all musically trained, is not a surprising result in the light of previously stated contentions.

[3] Seashore's method of grading a subject (6, p. 156; 1, p. 146) is, apparently, based on the assumption that each pair presents the same degree of difficulty in judgment. A subject's rank, therefore, simply depends upon the *number* of correct judgments irrespective of the distribution in the test of these correct judgments. Thus, preference of a minor second to an octave is no worse an error than preference of a perfect fifth to an octave.

[4] Percentile rank according to Seashore.

General increase in the number of errors found in the repetition may be due to a number of causes. Fatigue may play a dominant rôle. If it is true that fatigue, general or specific, follows repetition of the test, then successive reactions to a series of records played consecutively cannot be measured by some constant which does not compensate for these bodily changes. It must be remembered that the human organism is an extremely complex mechanism in continual flux, and that general lag and, especially, momentary fluctuations of attention due to any number of physiological changes may produce widely different test results. This is borne out in such a type of response as found in pair no. 13 for both tests. In the first test, 7 subjects preferred the minor second to the octave: in the second test, 13 subjects preferred the minor second to the octave. Only 4 subjects consistently preferred the minor second to the octave in both tests. Such an unusual type of response may find an explanation in fluctuation of attention. It is an unfortunate circumstance that the nature of the test under consideration does not permit any direct method for objectively detecting attention-fluctuations.

However, when fatigue actually enters into one's reaction, cognitive discrimination becomes the more difficult to effect, and there is a general tendency on the part of the organism to wander from the demands of the instructions and to react according to momentary mood. In a condition where concentration and rationalization have become minimized through fatigue, there is reason to expect that reaction to the tension-repose dimension of feeling, produced by various harmonic relationships, becomes more intensified, and, consequently, more frequent. In fact, such is found to be the case. If a study of judgment-reversals [5] actually indicates in an approximate manner what it is intended to indicate, then reaction to harmonic sequence is more liable to occur in a test repetition. Of the 35 subjects tested, 23 subjects made more reversals on the second test than on the first test: 4 subjects made the same number of reversals on both tests (however,

[5] See Table II., showing distribution of judgment-types.

differently distributed): 8 subjects made fewer reversals on the second test. It must be recalled that these reversals represent errors.

Let us briefly consider the case of two female subjects who took the test at a different sitting. Subject *A* has had 14 consecutive years of musical training. At a well-recognized musical conservatory, she studied the piano under one of the best masters in this country. Subject *B* has also pursued conservatory study and has had 10 years of musical training. In Seashore's consonance test, subject *B* received a percentile rank of 78, while rank of subject *A* was but 36. Ten of the 17 errors which subject *A* made were in the nature of reversals (that is, when she preferred the correct interval in a particular pair, she preferred the incorrect interval in the same pair reversed). In the case of subject *B*, 11 of the 13 errors which she made were in the nature of reversals. Upon questioning the subjects after the test, they declared that factors of resolution and harmonic progression disturbed them considerably in spite of every effort which they exerted to ignore such factors.

Three female subjects [6] possessing a background of wide musical training and experience were employed in a test designed for the purpose of subjectively ascertaining loudness differences of tones which might occur within any single binary combination as recorded on Seashore's phonograph disc. Each subject was presented with a blank form upon which she listed her judgments. The number of judgment agreements on such loudness differences was surprisingly large. In fact, only a few of the total number of intervals were found to be of uniform loudness in respect to both tones. In pair no. 10, for example, the upper tone of the first interval was found to be emphatically louder than the lower tone, thus tending to subdue the lower tone and produce greater unison. Especial mention is made of this particular interval, since this marked loudness difference affords a sufficient explanation why so many subjects preferred this interval

[6] These three subjects, *X*, *Y*, and *Z*, respectively have had 5, 10, and 11 years of conservatory training. All three have appeared in conservatory recitals and public performances. All three subjects play the piano. *Y* and *Z* are music teachers.

(minor second) to the one following (major second). In the major second, the roughness of beats is decidedly more marked, partly because of the greater uniformity of intensity of the two individual tones. In pair no. 31 (in which pair the intervals of pair no. 10 are in reverse order), the roughness of the minor second is very pronounced: this can be partly attributed to greater uniformity in loudness of the two tones. Indeed, the judgments rendered by the first group of 35 subjects seem to verify this. No subject, in either test, preferred the minor second in pair no. 31. However, in pair no. 10, 19 subjects in the first test and 18 subjects in the second test preferred the minor second to the major second:

TABLE II (*Group 1*)

TABLE SHOWING DISTRIBUTION OF JUDGMENT-TYPES

Intervals by Pairs	Type I			Type II			Type III			Type VIII		
	1	2	B	1	2	B	1	2	B	1	2	B
1–40	4	5	9	—	—	—	—	—	—	5	4	17
2–39	4	5	8	1	2	—	3	1	6	5	5	8
6–35	6	9	9	—	—	1	5	3	5	5	4	4
8–33	3	5	1	5	3	3	1	2	1	5	4	16
9–32	4	5	2	—	—	—	—	—	—	5	4	24
11–30	2	2	1	7	6	11	2	3	2	4	4	6
12–29	3	5	—	1	—	2	—	—	—	5	4	24
14–27	1	7	19	—	—	—	—	—	—	7	1	8
15–26	1	1	—	—	—	—	—	—	—	1	1	33
19–47	2	4	2	—	2	1	2	1	—	5	2	23
20–46	6	7	3	6	6	3	5	4	2	7	7	3
22–44	—	6	5	4	6	1	—	—	—	11	3	14
24–42	2	6	—	4	4	—	1	2	—	8	3	20

Intervals by Pairs	Type IV			Type V			Type VI			Type VII		
	1	2	B	1	2	B	1	2	B	1	2	B
3–38	6	1	—	5	6	12	1	1	—	4	8	7
4–37	3	1	—	1	8	19	1	3	2	7	—	2
5–36	6	4	8	4	6	2	4	5	1	5	4	5
7–34	9	3	2	5	7	5	5	9	2	6	6	1
10–31	7	6	12	—	—	—	—	—	—	6	7	10
16–50	6	5	9	2	6	—	6	6	—	7	4	5
17–49	1	5	1	7	6	1	3	4	—	8	4	14
18–48	5	2	1	1	2	—	1	—	—	4	7	23
21–45	6	10	2	5	5	1	4	7	2	9	2	6
23–43	4	2	—	7	11	6	3	4	2	10	7	3
25–41	7	9	6	1	2	3	6	8	5	7	2	—

12 subjects consistently preferred the minor second in both tests. (See Table II.) The rise and fall of intensity occurring within a number of binary combinations seemed to be another disturbing element to the 3 subjects judging loudness differences.

Explanation of Table II

Table II. indicates the distribution of judgment-types for all pairs. The pairs are numbered according to their sequential order in the test and are grouped in twos, according to reversal of the same interval content. (Grouping the same as in Table I.) The judgment-types indicated in Roman numerals are identical with the judgment-types classified under Roman numerals previously in the text. Under each type in the table are the figures—*1 2 B*. The column of figures under '1' represents the number of subjects who made that particular judgment-type in the first test but not in the repetition. The column of figures under '2' represents the number of subjects who made that particular judgment-type in the repetition but not in the first test. The column of figures under 'B' represents the number of subjects who made that particular judgment-type in both the first test and the repetition.

Thus, by way of example, in pairs 12–29 (see Table I.) we have in Table II. the possibility of four judgment-types. (For explanation of these types, see text.) Under type I., we have 3 subjects who made that particular judgment-type in the first test but not in the second test or repetition: 5 subjects made the same judgment-type in the second test but not in the first test. The dash under 'B' means that there were no subjects who consistently made the same judgment-type in both tests.

Again under type VIII. we have 5 subjects who made that particular judgment-type in the first test but not in the second test: 4 subjects made the same judgment-type in the second test but not in the first test: 24 subjects made the same judgment-type in both tests. To find the total number of subjects that made for a particular grouping of pairs a particular judgment-type in the first test, add the number of subjects

appearing under columns '1' and 'B': to find the total number of subjects that made the judgment-type in the second test, add the number of subjects appearing under columns '2' and 'B.' Thus under type VIII. for pairs 12–29, 5 subjects appear under '1,' 4 subjects appear under '2,' and 24 subjects appear under 'B.' The total number of subjects who made judgment-type VIII. in the first test is '1' plus 'B' or 5 plus 24 = 29. The total number of subjects who made judgment-type VIII. in the second test is '2' plus 'B' or 4 plus 24 = 28.

Table II. needs to be studied in conjunction with Table I. and the exposition in the text of the various judgment-types.

TABLE III (*Group 1*)

TABLE SHOWING PER CENT. OF CORRECT JUDGMENTS AND PERCENTILE RANK OF EACH SUBJECT FOR THE CONSONANCE TEST AND ITS IMMEDIATE REPETITION

Sub.	Test I		Test II		Sub.	Test I		Test II	
	% Right	Rank	% Right	Rank		% Right	Rank	% Right	Rank
1	62	18	64	26	19	66	36	72	68
2	62	18	58	8	*20	82	99	72	68
3	60	12	60	12	21	68	46	56	6
4	78	93	68	46	22	74	78	68	46
* 5	72	68	62	18	23	68	46	60	12
* 6	68	46	74	78	24	74	78	62	18
7	78	93	76	86	25	74	78	72	68
8	64	26	56	6	26	66	36	68	46
9	76	86	72	68	27	78	93	74	78
*10	60	12	54	4(–)	28	66	36	62	18
11	66	36	68	46	29	52	4(–)	62	18
*12	70	56	60	12	30	66	36	72	68
*13	78	93	78	93	31	76	86	66	36
14	70	56	64	26	32	70	56	62	18
15	66	36	56	6	33	72	68	64	26
16	66	36	66	36	34	56	6	62	18
17	72	68	64	26	*35	78	93	68	46
18	74	78	66	36					

N.B. An asterisk before a subject means that the subject, according to his report in the questionnaire, has had academic musical training extending over a number of years.

Although such observations may lack quantitative precision, their general agreement would reasonably lead one to believe that some marked variations in intensity actually do exist. Such variations could readily be demonstrated ob-

jectively through proper experimental means; however, it was not the writer's purpose to enter into such an experimental analysis of the intensity factor, but merely to indicate the possible existence of intensity variations which might have occurred as a result of one or more of the following conditions: first, the source of these variations may have been in the actual production of these tones; secondly, the source of the variations may have been introduced in the recording process of the initial production; and, thirdly, there may have been some defect in the phonographic disc employed. The latter case seems very unlikely in view of the homogeneity of the

TABLE IV (*Group 2*)

TABLE SHOWING DISTRIBUTION OF JUDGMENT-TYPES

(For explanation, see explanation under Table II.)

Intervals by Pairs	Type I			Type II			Type III			Type VIII		
	1	2	B	1	2	B	1	2	B	1	2	B
1–40	1	4	2	1	—	1	2	1	—	4	3	19
2–39	4	5	1	4	5	4	7	1	—	1	5	9
6–35	4	9	3	3	2	—	5	6	6	7	2	2
8–33	6	2	1	3	5	3	3	7	2	5	3	7
9–32	4	1	4	2	2	—	3	2	—	1	5	16
11–30	4	1	1	9	10	2	4	4	2	4	6	4
12–29	2	1	1	8	—	1	1	1	—	1	10	16
14–27	7	5	4	1	—	—	3	3	—	4	7	11
15–26	—	1	—	—	—	1	1	2	1	2	—	25
19–47	5	4	5	3	1	1	—	3	—	6	6	10
20–46	4	6	2	5	6	1	3	2	2	7	5	6
22–44	2	8	3	7	—	1	6	—	1	1	8	9
24–42	1	4	1	3	3	1	2	1	—	5	3	17

Intervals by Pairs	Type IV			Type V			Type VI			Type VII		
	1	2	B	1	2	B	1	2	B	1	2	B
3–38	—	—	—	3	4	14	3	1	1	3	4	6
4–37	—	4	—	8	5	11	4	3	3	4	4	—
5–36	2	5	2	6	8	1	8	5	—	7	5	4
7–34	3	3	2	7	7	5	6	5	1	4	5	2
10–31	4	4	3	1	—	—	—	2	1	4	3	17
16–50	4	7	8	3	3	1	1	5	1	9	2	3
17–49	2	1	—	7	8	—	1	3	—	9	7	11
18–48	2	—	—	3	5	—	1	3	1	8	6	15
21–45	2	6	2	2	4	1	3	2	4	8	3	8
23–43	4	3	1	6	6	7	4	4	—	5	6	3
25–41	4	4	6	1	3	2	7	3	5	2	4	3

disc material and the considerable care and attention given to the manufacture of phonographic records. Besides, in the particular experiments under consideration, much care was exercised in the handling of the record and the use of new needles. Although not at all impossible, the second case seems quite as unlikely as the third.

It is in the first case cited that the probability of introducing variations in intensity seems greatest. Indeed, whenever pipes or reeds are employed as a source of sound, any slight difference in air pressure produces variations in intensity. The difficulty of producing staccato effects and of avoiding any such variations in intensity is quite obvious in view of the sluggish action of most valves in controlling the source-volume of air admitted into the various reeds or pipes as the case may be. Any such sluggish action would cause a gradual rise and fall in intensity of the sound emitted. If it is to be assumed that the variations noted find their source in the original production, then there follows an additional difficulty far greater in significance than merely any deficient control of the intensity factor *per se*. It is a well known fact that in the case of wind instruments, variations in intensity, *i.e.*, variations in air pressure, actually produce appreciable pitch differences. If such a condition of affairs existed in the original production, then one is immediately led to feel the decided uncertainty of the test material, since it is quite impossible to evaluate the actual pitch relations produced under such variable conditions. If variations in intensity are introduced, then, by way of example, an octave no longer represents a pure consonance, but becomes, in reality, a marked dissonance.

However the conditions may have been at the time of the original production, let it be assumed that the intensity factor was so carefully controlled that it represents a constant throughout the entire series of binary combinations. Is there any other possible explanation for variations in loudness which have been found to exist in reproduction? Emphatically yes. The resonance chamber of the phonograph plays a most significant rôle in the amplification of certain tones. It is

safe to say that no two resonance chambers of any two phonographs, whether they be of the same design and make, or of different makes, respond to sounds in identically the same fashion. In some phonographic chambers definite pitch relations of certain tonal qualities may be well defined, whereas in other chambers the same definite tonal relations may be markedly ill-defined. Some chambers tend to amplify certain partials, while other chambers do not provide any such amplification. Indeed, any attempt to match with scientific accuracy the resonating chambers of any two phonographs would represent an exceedingly difficult undertaking. This fact leads one to question the validity and the accuracy of results obtained from the use of different phonographs when compared to a norm established on the results obtained from the use of a particular instrument possessing definite characteristics.

Position of the phonograph in respect to the subjects, and the size of the room in which the test is given, are by no means insignificant factors. It is reasonable to expect different reactions from subjects seated from the phonograph at different distances and at different angles. Comparison of results of various investigators again becomes an uncertain procedure in view of Seashore's allowance of needles differing in their production of loudness to be employed at the discretion of the experimenter.

III

On April 21, 1925, at 5:00 P.M., the Seashore consonance test was given to a group of 30 subjects representing a class in General Introductory Psychology in the Teachers College at the Johns Hopkins University. This group (Group 2) was made up of 17 women and 13 men. It was a most fortunate circumstance that of these 30 subjects, 15 subjects had had appreciable musical training and experience. (See Table V.)

The test was given under the same general conditions as that given to the 35 subjects (or Group 1) whose results have already been discussed. However, a few slight variations in test procedure should be noted at this time. Unlike before, the questionnaires were not distributed before the

test had begun. Instead, the first sheet of the questionnaire (which includes three printed sheets of questions) was distributed among the subjects during the intermission or rest-period following the first test and preceding its repetition. The rest-period approximated the same length of time as that granted to the subjects in the first group. The remaining two pages of the questionnaire were distributed among and completed by the subjects at the close of the repetition of the test.

TABLE V (*Group 2*)

TABLE SHOWING PER CENT. OF CORRECT JUDGMENTS AND PERCENTILE RANK OF EACH SUBJECT FOR THE CONSONANCE TEST AND ITS IMMEDIATE REPETITION

Sub.	Test I		Test II		Sub.	Test I		Test II	
	% Right	Rank	% Right	Rank		% Right	Rank	% Right	Rank
* 1	74	78	64	26	*16	70	56	68	46
2	34	4−	34	4−	*17	76	86	78	93
* 3	66	36	68	46	18	58	8	68	46
* 4	70	56	66	36	*19	58	8	82	99
5	66	36	68	46	20	64	26	54	4−
6	68	46	42	4−	*21	60	12	72	68
7	68	46	78	93	22	54	4−	50	4−
* 8	64	26	72	68	23	62	18	64	26
* 9	66	36	70	56	24	78	93	80	97
10	60	12	50	4−	*25	70	56	66	36
*11	62	18	70	56	26	52	4−	56	6
12	76	86	72	68	27	68	46	70	56
13	62	18	52	4−	*28	74	78	76	86
*14	80	97	76	86	29	62	18	60	12
*15	70	56	70	56	*30	66	36	82	99

N.B. An asterisk before a subject means that the subject, according to his report in the questionnaire, has had academic musical training extending over a number of years.

Female subjects are in italics.

Subjects 14 and 17, because of previous familiarity with Seashore's consonance test, cannot rightfully be compared with the other subjects in this group.

Before the phonograph was played, the following instructions were read to the listeners:

You have been given a slip of paper upon which are printed numbers arranged in serial order from one to fifty. Upon the horizontal line after each number, you are to write a certain judgment of your own in reference to something which you will hear on this phonograph. That means that you will make a total of fifty judgments.

You will hear two combinations of two tones each: one combination is better or worse than the other in consonance (harmony). A good combination is one in which

the two tones are smooth, and blend, tending to fuse together into one. A bad combination is just the opposite. If the second combination is better, record B; if Worse, W.

Fifty pairs of such combinations will be played for you on the phonograph. No two combinations are the same in any pair: the second combination in any particular pair is either *better* or *worse* than the first combination. If you are in doubt whether the second combination of a pair is better or worse than the first combination, then *guess*.

Do not converse with your neighbor at any time during the test, nor at the end of the test: it is understood, of course, that you are not to look upon your neighbor's paper at any time.

After these instructions had been read by the director, the subjects were allowed to ask questions relating to a correct understanding of what had been read. A couple of simple questions were raised in reference to sentence content: these questions were immediately answered. Upon answering these questions, the director again repeated the instructions in full before playing the record. Before the test was repeated after the intermission, the directions were read a third time.

The results of this experiment were strikingly similar to the results obtained from the first group. Reversal in judgment was quite marked, especially in those pairs in which factors of immediate harmonic progression can be held to account for such reversal. By way of comparison with the pairs previously discussed in the text under the first group, in the second group of 30 subjects, 13 subjects in the first test and 13 subjects in the repetition preferred the major third in pair no. 23 and the perfect fourth in pair no. 43. Seven subjects consistently made the same reversal in both tests.

In pair no. 3, in the first test, 17 subjects preferred the octave to the perfect fifth, while in pair no. 38, the same 17 subjects preferred the perfect fifth to the octave. In the repetition, 18 subjects made this same type of reversal. (Compare Tables II. and IV.) Fourteen subjects consistently made the same reversal in both tests.

Again, in pair no. 4, in the first test, 19 subjects preferred the minor sixth to the minor third, while in pair no. 37, the same 19 subjects preferred the minor third to the minor sixth. In the repetition, 16 subjects made the same type of reversal. Eleven subjects consistently made the same reversal in both tests.

Again there occurs preference for the relatively rich combinations. In pair no. 28, in the first test, 16 subjects preferred the major third to the octave, while in the repetition, 15 subjects preferred the major third to the octave. Thirteen subjects consistently made such a preference in both tests. (The reader must not lose sight of the fact that in the first group of subjects, five more subjects were tested than in the second group.)

Unlike the results obtained from the first group, the second group evidenced a slight general improvement in the repetition. Of the 30 subjects tested, 28 subjects scored a different grade on the first test than on the repetition. Of these 28 subjects, 12 subjects showed an increase of errors in the repetition, while 16 subjects showed a decrease of errors in the repetition. Two subjects made the same number of errors in each test: however, again in none of the individual records was the error distribution the same for both tests. In view of individual differences, this slight improvement shown by the second group does not invalidate the previous contention that fatigue, specific or general, may play a prominent rôle in the repetition of the test. In the matter of reversal of judgment, in the second group 11 subjects made fewer reversals on the repetition, 14 subjects made more reversals on the repetition, while 5 subjects made the same number of reversals in both tests (but differently distributed for each individual in both tests).

The same inconstancy of judgment between the first test and its repetition observed in the results of the first group was largely duplicated in the results of the second group. (Compare Tables II. and IV., III. and V.) In Table V., notice especially the extraordinary jump in percentile rank for the second test of subjects 7, 19, and 30. Also notice that, in general, the musically trained subjects scored relatively low.

At the close of the second test, the following additional written information was obtained from the subjects in response to the following list of questions:

(1) As far as you know, did any elements of harmonic progression enter into your judgments?

In answer to this question, 14 subjects answered "yes," 9 subjects answered "no," 3 subjects answered that they did not understand, one subject answered that she did not think so, while another subject answered that he did not know.

(2) Can you recall the musical structure of any particular pair of combinations in which your judgment was so affected?

Subject 9 was able to accomplish this. She referred to a pair which contained an octave and a third (obviously pair no. 28). Study of her record for pair no. 28 revealed the fact that in each test she had preferred the major third to the octave.

In answer to this same question, subject 24 replied: "I cannot just explain this, but several combinations were from some familiar pieces." This reply again points out how association may function in listening to a series of intervals.

Subject 28 replied that "Several needed by one more chord in order to resolve the tone. A chromatic drop of the base was noticeable." Assuredly, this reply convinces one that this subject was reacting to various elements of harmonic progression.

(3) Have you ever had a course in harmony?

Subjects 4, 11, and 25 answered "yes."

(4) Have you ever taken a course in physics?

Eighteen subjects answered "yes."

(5) Did this course include the physics of sound?

Sixteen of the eighteen subjects who answered "yes" to question no. 4, answered "yes" to this question.

(6) Have you had a course in acoustics?

Two subjects answered "yes."

(7) Can you recognize intervals?

Nine subjects answered "yes." These subjects are as follows: 1, 3, 4, 9, 13, 19, 21, 25, 26.

(8) Give the interval value of any interval which occurred in the test.

Subjects 4, 9, 21, and 25 recorded certain interval values which had occurred in the test.

Questions 2 to 8 inclusive have been considered of much importance in a study of judgment-types. If a subject is able to recognize intervals, and if his tendency is to attach a certain consonance value to any particular interval which he may hear, then upon repetition of any particular interval, he is likely to attach the same consonance value to that interval. Such practice, of course, largely depends upon the subject's ability to recognize intervals in any musical environment, and to remain indifferent to the various elements of harmonic progression. In the case of a subject effecting such recognition and applying a preëstablished value to such recognition, we can readily expect that when his preëstablished value is incorrect, then his judgment of a repeated interval is bound to be incorrect. It was this fact that led to a study of judgment-types. For example, in pair no. 37, a total of 26 out of 30 subjects preferred the minor third to the minor sixth. However, of these 26 subjects, 7 also preferred the minor third in pair no. 4. That means that 7 subjects preferred the minor third irrespective of its position in the pair. Although one cannot be positively certain, it would seem that these seven subjects were applying in pair no. 37 a preëstablished consonance value to an interval which they happened to recognize. Although the case may be a doubtful one, the writer has given these subjects the benefit of the doubt in assuming that in making such a judgment-type (that is, consistently erring in a pair and its reversal), the subject is not necessarily being influenced by the harmonic relation between intervals. If this distinction were not made, then still more subjects could be added to the list in substantiation of the fact that harmonic progression does affect judgment.

IV

In conclusion, what can be said regarding the validity of the test for revealing a so-called 'sense' of consonance? From the foregoing analysis, one is inevitably led to conclude that *paired interval comparison is an inadequate method for testing consonance*, since, as evidenced, such comparison is conducive to reaction to the various elements of musical

progression, including the feelings which, apparently, are inseparably attached thereto. Assuming that the musically talented have a 'sense' of consonance, and since the object of this test is to determine and point out such musical talent, from the very nature and structure of the test material, there is reason to expect negative results from the talented group.

All of which raises a very pertinent question—a question which needs to be considered in the light of musical history and of modern musical tendencies, and which must be left unanswered until more extensive and intensive investigations have been carried out; namely, is a sense of consonance [7] a fundamental and necessary factor of the 'musical mind'?

REFERENCES

1. GAW, E. A. A Revision of the Consonance Test. Psychol. Monog., 1918, XXV., No. 2, 134–147.
2. HELMHOLTZ, H. L. F. On the Sensations of Tone. Trans. by Ellis. London: Longmans, Green, and Co., 4th ed., 1912.
3. MALMBERG, C. F. The Perception of Consonance and Dissonance. Psychol. Monog., 1918, XXV., No. 2, 93–133.
4. PETERSON, J. A Functional View of Consonance. *Psychol. Rev.*, 1925, XXXII., 17–33.
5. POLE, W. The Philosophy of Music. London: Kegan Paul, Trench, Trubner and Co., Ltd., 1910, 5th ed.
6. SEASHORE, C. E. The Psychology of Musical Talent. New York: Silver, Burdett, and Co., 1919.
7. SEASHORE, C. E. The Measurement of Pitch Discrimination: A Preliminary Report. Psychol. Monog., 1910, XIII., No. 1, p. 41.
8. SEASHORE, C. E. Manual of Instructions and Interpretations for Measures of Musical Talent. New York: Columbia Graphophone Co., 1919.

[7] That is, what Seashore and his collaborators mean by the term 'consonance.'

THE SPATIAL THRESHOLD OF TOUCH IN BLIND AND IN SEEING CHILDREN[1]

BY MARGARET S. BROWN AND GEORGE M. STRATTON

There have been diverse reports by experimenters upon the tactile sensitivity of the blind. An account of the earlier work will be found in the elaborate study by Griesbach,[2] and need not here be repeated, save to say that while some, notably Czermak, Gärtner, and A. Stern, had found the blind to be superior to the seeing in the tactile discrimination of space—the 'two-point threshold'—Uhthoff had found no difference between the two kinds of individuals. Griesbach's own experiments, on the contrary, led him to believe not only that there was a difference, but that, slight though it was, it was to the disadvantage of the blind. Some experiments by Miss Bunnell and Miss Stowell, reported by one of the present writers,[3] had supported the earlier position that the blind were slightly superior. Finally in this field of tactile space Seashore and Ling[4] found that the threshold for the blind was neither lower nor higher than for the seeing, although the seeing had a somewhat greater variability than had the blind.

The reason for such diversity of finding perhaps lies in the fact that the conditions under which the experiments were conducted were by no means uniform. The kind of æsthesiometer used, the amount of pressure exerted upon the skin, the fatigue, the degree of attention given, the chance suggestions received—all these are important modifying factors.

[1] From the Psychological Laboratory of the University of California. The writers are indebted to Dr. Warner Brown and to Dr. R. S. French for invaluable assistance in this study.

[2] H. Griesbach, 'Vergleichende Untersuchungen über die Sinnesschärfe Blinder and Sehender.' *Archiv. f. d. gesam. Physiol.*, 1899, Vol. 74, pp. 577–638; Vol. 75, pp. 365–429, 523–573.

[3] Stratton, 'Experimental Psychology,' 1903, p. 46.

[4] C. E. Seashore and T. L. Ling, 'The Comparative Sensitiveness of Blind and Seeing Persons.' Psych. Mon., Vol. 25, 148–158.

Only in the work of Seashore and Ling was there care that the seeing and the blind should be of approximately equal intelligence; and yet even here the age was not kept fairly equal, for the blind were older than the seeing. And in none of the work does there seem to have been any attention given to the consistency of the findings—the 'reliability,' so-called, of the tests. It is therefore clear that in the field of space-perception by the blind there is both opportunity and need of further work.

I. The Method

In planning the experiment, an effort was made to avoid certain sources of error which are easily present in testing for the two-point and pressure thresholds by the ordinary method. Of these errors, suggestion, and the strain involved in waiting for the contrasts are probably the worst. They seemed particularly undesirable with subjects who are already handicapped in sensory equipment. The chief effort, therefore, was to find a means of testing for these thresholds which would do away with suggestion, and which would lessen the strain of expectation. If it were possible to have the blind work in a way to which they were accustomed, it would decrease these undesirable factors.

To accomplish this end a method was adopted which permitted the child to choose his own time and rate of touching the points; a method which used, instead of trying to inhibit, the tendency to move the fingers when making fine tactual discrimination; and which, by the essentials of the method of serial groups [1] reduced greatly the influence of suggestion. An æsthesiometer was made, having many raised points, over which the subject ran his finger, as over a page of Braille. The points were arranged in rows, the rows an inch apart; and in the rows there were irregular alternations of single points and of pairs of points, the two points of each pair being set at varying but accurately measured distances apart. In any given row there were as many single points as there were pairs. And these were spaced, also an inch apart, so that

[1] Stratton, 'The Method of Serial Groups,' *Psych. Rev.*, IX., 444.

there were rows along the board either lengthwise or crosswise. As the subject moved his finger along from left to right, the distance between the two points of such pairs as occurred in the row decreased, discrimination growing increasingly difficult. These points were, in reality, tiny steel posts raised 3 mm. uniformly above a solid smooth surface, each post being .34 mm. in diameter, made of a No. 12 Milward needle cut off squarely at the top and polished. These were set with precision by an expert instrument maker, the distances being computed from the centers of the posts. There were, in all, eighteen pairs of points, forming a series decreasing from the most distant, which were 2 mm. between the points, to the least distant, which were 0.3 mm. apart, the decrease being by 0.1 mm. as one passed from pair to pair.

It was desirable that all the subjects put forth their best efforts throughout the tests; and to accomplish this end, it was decided that competition should be introduced, by offering prizes for the best scores. Three prizes were given for the best scores on the first sitting, and three for the second sitting. The children were tested individually but the competition was within each of several groups: of the blind children, those of the fourth, fifth, and sixth grade forming one group, and those of the seventh, eighth, and ninth grade forming another; the control subjects, the seeing children, formed the third group. Thus each group received a prize.

The blind children tested were thirty-five in number, from the California State School for the Blind. These were all the available children in the regular classes from the fourth through the ninth grade, between the ages of twelve and eighteen. Intelligence scores were available for twenty-four of the thirty-five blind subjects. By this means, and since only those children were taken who were doing regular work, it was possible to eliminate those of low intelligence, such as might be found in the special classes of the school.

The blind children thus selected were next classified into five degrees of sight which were determined by simple tests, with care always for intense and equal illumination.

Degree 1, was total loss of vision. The organ of sight was

frequently lacking in these subjects. No light whatever could be seen.

Degree 2 meant the power to distinguish between light and darkness, but not to recognize the form of objects. Strong sunlight reflected directly into the subject's face from a sheet of glazed white cardboard elicited the response 'lighter,' while exposure of a black sheet in the same position brought forth the response 'darker.' In no case was the subject able to perceive the shape of the object used as a reflector.

Degree 3 meant only the coarsest vision for the form of objects. A sheet of glazed white cardboard 56 cm. square was used, with a band 20 cm. wide and 56 cm. long extending across the middle of the sheet. The cardboard was held vertically 1 meter from the subject's face, and the subject was required to judge whether the black band extended vertically or horizontally across the white sheet. Out of three trials, all judgments were required to be correct.

Degree 4 required that the subject be able to describe with general correctness the appearance of a 56-cm.-square sheet of paper held 1 meter away, the face of which was divided into halves, the one half being entirely of a uniform gray; and the other half, bearing a checker-board design of alternate black and white squares, each square with sides of 5 cm.

Degree 5 required that the subject be able, by sight, to distinguish large letters. Printed letters 1¼ inches in height were used. The subject was required to name the letters, or read the word before him. Some of these subjects could read normal print held close to the eyes but none could read it at a distance of 14 inches. The children in this group were able to read Braille characters, not only by touch, but visually. Some of these children, upon occasion, bring their eyes close to the page of Braille, as if visually to supplement their tactile perception of it.

Since the blind children were under institutional care, it was deemed advisable to choose the seeing children likewise from an institution. Accordingly, for comparison, fifteen girls from the Roman Catholic Orphan Asylum in San Francisco were selected and these girls were between the ages

of twelve and seventeen, and between the fifth and the tenth grades, inclusive. Intelligence scores in their case were not available, but no girl was chosen who was not making normal scholastic progress, or who gave evidence of maladjustment or incorrigibility. From these normal children, the experimental group of fifteen was selected at random by the Sister in charge, with age requirements chiefly in mind. The two groups were therefore approximately equal in age, school, and living conditions.

All tests were given by the same experimenter. The blind children were tested between 8:00 A.M. and 1:00 P.M. The seeing children were tested between 9:00 A.M. and 4:00 P.M. Each subject had three sittings: a preliminary sitting, in which he became accustomed to the apparatus, the procedure, and the experimenter, and was told something of the purpose of the test. Record was made only of the next two sittings, which were always at least four days apart.

The subjects, as was said, were tested individually. They were permitted to use the index finger of the preferred hand. Of the blind, this was the right hand in thirty-four out of the thirty-five cases. Of the seeing, it was always the right hand. The seeing children were blindfolded for both sittings. The blind children were never blindfolded, but were placed so that the instrument was out of their range of vision.

The directions to the subjects were as follows: "I have here a board in which are placed some little steel posts. They are set in two long rows down the board. They also form short rows across the board. Sometimes there is one post, and sometimes there are two near together, but never more than two. I want you to tell me how many you feel in position—whether one or two. Take them across this way first." The subject was required to take the rows across the board on the odd runs down the board, and to take the rows longitudinally on the even runs. This prevented his becoming familiar with the position of the pairs. Portions of the board which a subject gave correctly several times were omitted for him in later runs; that is, if he always recognized the upper five pairs correctly, he was no longer required to

give judgment upon the upper two or three of these pairs. If he always was in error below a certain position, then the lower of these pairs were dropped out.

By this procedure, each sitting yielded from each subject ten judgments for each separation of the two points. The threshold was arbitrarily fixed as the least separation at which 80 per cent. of the judgments for any given subject were correct. Provision was made for exceptions to this rule. Occasionally errors were made in judging points much farther apart than those at which the errors began to appear consistently. If no one of the three degrees of separation next lower than this single case gave more than two errors, then the errors made in judging the greater degree of separation were disregarded. An example will make this clear. One observer made no error in judging points 2.0 mm. apart, and no error with points 1.9 mm. apart; at 1.8 he made four errors; but in the series of distances from 1.7 to 1.0 he in no case made more than two errors. At 0.9 mm. he made four errors; at 0.8 mm. four errors; at 0.7 mm., seven errors; and three or more errors at each smaller distance than this. His threshold was set down as at 1.0 mm. and not as at 1.8 mm., which a strict application of the rule would have required.

The blind subjects were found, by the tests indicated, to be distributed as follows, among the several degrees of sight.

	Degrees of Sight				
	I	II	III	IV	V
Boys	6	3	0	1	6
Girls	4	6	1	0	8
Total	10	9	1	1	14

In the comparison to be reported hereinafter, between the tactual space of those of different degrees of blindness, the degrees III. and IV. are disregarded because the number of individuals in them is so small. The 19 subjects in degrees I. and II. who can see no shape whatever were regarded as

totally blind, while the 14 subjects in degree V. were regarded as partially blind.

From these two groups, in order to bring out possible differences, separate statistical attention was given to those totally blind from birth, of whom there are eight, and those partially blind from birth, of whom there are eleven.

II. The Results

To obtain the coefficient of reliability of the test itself the scores obtained for each subject in the first sitting were compared with the scores of these same individuals in the second sitting.

The coefficient of correlation between the two sittings indicated that the test gave a reliable measure of the spatial threshold.

With the seeing group, there was a lower coefficient of reliability than with the blind. This may be due to the fact that the distances at the upper end of the æsthesiometer are too small to be an adequate measure for seeing children. The widest distance between points, it will be remembered, is 2.0 mm.

1. Two-Point Thresholds in Millimeters

Returns from the Two Sittings Given Separately

I. = First Sitting; II. = Second Sitting

Group	No. Cases	Mean I.	Mean II.	S.D. I.	S.D. II.	Correlation of Means I. and II.
Whole group of blind	35	1.20±.05	1.17±.06	.40	.50	+.76±.05
Seeing group	15	1.77±.05	1.70±.06	.26	.32	+.54±.12
Partially blind	14	1.36±.05	1.39±.08	.29	.44	+.62±.11
Totally blind	19	1.08±.08	0.97±.07	.37	.45	+.79±.06
Partially blind from birth	11	1.36±.06	1.37±.08	.27	.38	+.46±.16
Totally blind from birth	8	0.91±.07	0.91±.11	.27	.45	+.58±.16

2. Two-Point Thresholds in Millimeters

Returns from the Two Sittings Combined

Group	No. Cases	Means and P.E. of Mean	S.D. of Distribution
Whole group of blind	35	1.160±.049	.435
Seeing group	15	1.780±.049	.285
Partially blind	14	1.364±.068	.378
Totally blind	19	0.973±.064	.413
Partially blind from birth	11	1.355±.077	.377
Totally blind from birth	8	0.925±.088	.369

3. Entire Group of Blind

Means and Correlations of Age, School-Grade, and Two-Point Threshold for Both Sittings Combined

Age		School-Grade		Two-Point Threshold	
Mean	S.D. of Distr.	Mean	S.D. of Distr.	Mean	S.D. of Distr.
15.11	1.83	6.43	2.03	1.16	.435

Coefficients of Correlation

Age with school-grade ..... +.758
Age with two-point threshold ..... −.001
School-grade with two-point threshold ..... −.002
Partial Correl.: School-grade with two-point threshold when age is rendered constant ..... −.001

Two-Point Thresholds: Difference of Means of Such Groups as are Mutually Exclusive

	Seeing	Partially Blind	Partially Blind from Birth	Totally Blind	Totally Blind from Birth
Seeing		.41±.07	.42±.11	.69±.11	.86±.12
Partially blind	.41±.07			.29±.08	
Partially blind from birth	.42±.11				.44±.13
Totally blind	.69±.11	.29±.08			
Totally blind from birth	.86±.12		.44±.13		

(Above is for First Sitting)

Seeing		.31±.14	.32±.14	.73±.13	.78±.18
Partially blind	.31±.14			.32±.16	
Partially blind from birth	.32±.14				.46±.20
Totally blind	.73±.13	.32±.16			
Totally blind from birth	.78±.18		.46±.20		

(Above is for Second Sitting)

	Seeing	Partially Blind	Partially Blind from Birth	Totally Blind	Totally Blind from Birth
Seeing		.416±.084	.425±.092	.807±.080	.855±.101
Partially blind	.416±.084			.391±.093	
Partially blind from birth	.425±.092				.430±.116
Totally blind	.807±.080	.391±.093			
Totally blind from birth	.855±.101		.430±.116		

(Above is for Both Sittings Combined)

III. Conclusions

1. The measurements obtained in the two sittings are correlated with each other to a degree which indicates that the method employed gives consistent, *i.e.*, 'reliable,' results.

2. The differences between the means of those groups which are mutually exclusive are sufficiently large in comparison with the probable errors of the differences to warrant the conclusion that the differences themselves are significant and are not merely accidental or specious.

3. The spatial threshold of touch for the blind group as a whole is lower than for the group of those who have sight. And this advantage over those who can see holds true for each of the sub-groups of the blind, namely for the partially blind, the totally blind, the partially blind from birth, and the totally blind from birth.

4. The degree of the blindness affects the spatial threshold. The totally blind have a lower threshold than the partially blind.

5. Among the blind, the threshold does not significantly vary with grade in school. The fineness of tactual perception of space appears to be an insignificant factor in the blind child's progress in school.

6. The present findings are in accord with those of the earlier experiments in California, and of Czermak, Gärtner and others, who found the blind to be superior to the seeing in the tactile discrimination of space; and are opposed to those of Uhthoff, Griesbach, Seashore and others who found either that there was no difference in this respect, between the blind and the seeing, or that the tactual advantage was actually on the side of those who could see.

THE DEVELOPMENT OF NUMBER CONCEPT IN CHILDREN OF PRE-SCHOOL AND KINDERGARTEN AGES.[1]

BY HARL R. DOUGLASS

University of Oregon

CONTENTS

1. Purpose and scope of the study 443
2. Methods and procedure 446
3. Number concepts of kindergarten children of ages 4½ to 6 years 454
4. Development of number concept with age 458
5. Sex differences 463
6. Validity and reliability of the tests 465
7. Summary and conclusions 466

1. PURPOSE AND SCOPE OF THE STUDY

It is the purpose of this discussion to throw some light on the problem of the development of the number concept in children before the age of school entrance; to bring to light data which may assist in arriving at a somewhat better idea of the degree of number concept possessed by children of pre-school ages and which may disclose variation among youngsters of the same age.

While the qualitative aspects of the development of number concept are indeed most inviting, this study concerns itself only with the quantitative side. Most interesting and valuable will be reports on the manner of the development of number concept as indeed will be discussions of the nature of number concept, when these contributions are made. In this study, the writer has confined himself to the less pretentious and less intriguing task of seeking data which may aid in gauging the extent to which children possess concepts of numbers at a series of pre-school ages—or to put it in another way, the quantitative aspect of the development of number concept in youngsters of the ages studied.

[1] The author is indebted to Professor Lewis M. Terman of Stanford University for suggesting the study, and to his wife, Zanna M. Douglass, for valuable assistance in the details of the experiment.

When one faces the problem, certain difficulties confront one immediately. In the first place, one asks oneself what constitutes the possession of number concept, say, of the number four. May the inclusion of the word four in the vocabulary be taken as evidence of such possession? Writers generally, including Major,[1] Arnett,[2] Rusk,[3] have made special note of the fact that youngsters may not only employ the word attached to the concept without possessing anything approaching the conventional concept but that even when counting involves the word, the concept may be, and in the early stages of counting usually is, lacking.

What then is the standard by which we shall judge possession of concept? This problem, encountered early in any quantitative study of concept possession or development, is even more prominent in the study of number concept because of the counting phenomenon. Is it necessary that a child recognize a group of four objects, let us say dots on a piece of paper, before he may be said to 'know' four? Investigations have shown that the child may 'know' four in some patterns of grouping and not in others,[4] that recognition with a given individual depends upon the arrangement of the objects, the proximity of them to one another and the color and size of the objects as well as upon the nature of the object itself. What plan of grouping, for example, shall be followed in a test for number concepts? Or should the standard be recognition of four in any and all possible groupings? Such a test can never be devised, as there is no limit to the number of possible different groupings. But does a child, even though able to recognize four objects regardless of plan of grouping so long as he is able to see them all without materially shifting the gaze, possess the ability to recognize four when

[1] Major (33), p. 163. "Further ability to point to objects one after another and want them is not in itself evidence of the knowledge of number for the reason that the counting may be imitative and mechanical."

[2] Arnett (2), p. 328. "It is only natural that a child in learning to count first masters the number names as a linguistic feat and only later comes to apply them in actual enumeration."

[3] Rusk (48), p. 295. "The ability to count or repeat numbers, say up to ten, only indicates that a series of closely associated terms have been learned."

[4] Freeman, Nanu, and Messenger particularly; see bibliography.

the four are, let us say, elephants of different sizes, or if the four be pieces of furniture, somewhat dissimilar? Must we insist that the 'knowing' of four depends upon perceiving four without counting? or adding? Does a child really 'know' four until he is able to assemble a group of four objects, to select four from a large number? Must he be able to distinguish four from three, from five, and all other numbers? Can he be said to have a 'true' concept of four if he is not aware of all of its properties, *e.g.*, that it is half of eight or a third of twelve, that it is twice two and the sum of three and one, and that it is the difference between ten and six and between five and nine? Must one not know that it is the square of two and the square root of sixteen, that it is the logarithm of 10,000? Can a child be said to possess the concept of the number four until he can identify tactile, auditory, kinæsthetic, or gustatory experiences of four in all the possible variations and situations? It is clear that there is no limit which may be set to the extension or perfection of concept. It is never complete, and the bounds of its development are limitless. Those who have approached concept studies have recognized this complication. What we find is varying degrees of "conceptness"—that is, varying degrees of attainment to a theoretically perfect concept. What one must do of course is to set standards, reasonable even if somewhat arbitrary, and the results of the particular study must be referred to this arbitrary standard for interpretation. In the experimental phase of this study which follows just this thing has been done.

The immaturity of the children studied aggravates the difficulty of the problem in this study. It is difficult to obtain standard conditions and because of the inability of pre-school youngsters to write, the study must be oral and individual, involving an expenditure of a large amount of time in ratio to the results obtained.

The specific questions upon which the data of the study have some bearing are:

1. Of what numbers from 1 to 10 does the kindergarten child (from ages four and one half years to six years) have

reasonably adequate concepts? In what degree does he have concepts of these numbers?

2. To what extent do youngsters of the same age vary in their possession of such concepts?

3. Are there sex differences in the possession and development of such concepts?

4. What is the nature of the growth curve representing the development of them?

5. What are the relative tendencies to overestimate and to underestimate the various numbers of objects from 1 to 10?

6. Is it possible to devise a means for evaluating varying degrees of concept possession as determined by the character of their inaccuracies of judgment?

2. Methods and Procedure

Three tests were given to from 26 to 40 pupils in two kindergarten classes of the public schools of Palo Alto, California. The ages of these children ranged from 4 years 6 months to 6 years. One of these tests was given three times, twice at the same testing and once a month later. The other two tests were given but once. The test which was given three times and which we will call the 'dot recognition' test, consisted in requiring the subject to estimate without counting the number of blue dots [1] on a card when they were exposed to easy view.[2] In cards containing dots [3] of numbers from one to ten were exposed. The white cards were 4 x 6 in. in dimension. The dots of a solid dark blue color were approximately ¼ in. in diameter and were arranged in a single row on the card with an intervening space between dots of approximately ¼ in.

[1] Visual number concept was selected for testing in the belief that this sense avenue would be most effective in stimulating the fullest expression, in most completely testing the possession of number concept.

[2] For abbreviated accounts of methods and materials of investigation of a similar or comparable nature see Howells (24), Ch. V., or original reports of investigations by Nanu, Lay, Freeman, Cattell, Warren, Burnett, Deitze, Messenger, references to which are given in the bibliography.

[3] In other investigations various types of stimulus objects have been used, usually figures on cardboard, dots, circles, squares, lines, other geometrical figures. Freeman used small disks of light projected on a screen.

The cards and dots were exposed just long enough for the child to see the dots and to contemplate the population.[1] Care was taken by the experimenter to withdraw the dots before they could be counted. This was done in the belief that the ability to determine the number of dots on a card by counting was not as satisfactory an evidence of concept possession as the ability to estimate the number of dots without counting, *e.g.*, a given child may be able to fix his attention successively upon each of the dots on the card, counting simultaneously in a mechanical fashion and concluding by announcing the last number 'counted' as the number of dots on the card and yet have little concept of that number. As a consequence of this possibility no standard time exposure was used, as such would obviously not meet the purposes of the experiment. Then, too, the impracticability of determining the proportional times for exposure of the different numbers of dots so that there would be equality of opportunity among the different numbers added to the undesirability of that plan of procedure. The writer is of the opinion that equality of opportunity among the children was as well, if not better, served by the method used than it would have been by employment of uniform exposure for all children. Equality of time is quite evidently not synonymous with equality of opportunity.

It may occur to one to question the choice of visual tests. It may be suggested that the possession of the number concept may also be tested by tests involving other senses, such as auditory or tactile impressions of number. In fact, it may be suggested that the degrees to which one possesses the concept of number is measured by the ability to recognize number through these senses as well as to recognize number through visual impressions. However, in the absence of any agreement as to just what does constitute number concept, it is reasonable to limit it to the abstract idea of exact manyness, independent of direct sense content. It seems sound to measure number concept by that sense avenue which is most

[1] Other investigators have almost uniformly used uniform exposures ranging from .01 second (Messenger, Cattell) to .13 second (Warren). In a more recent investigation (Freeman) the exposure was .05 second.

highly trained and sensitive, *i.e.* that one in which the lack of sense training is most unlikely to operate to prevent a most complete and favorable function of the number concept. For this reason visual tests were employed.

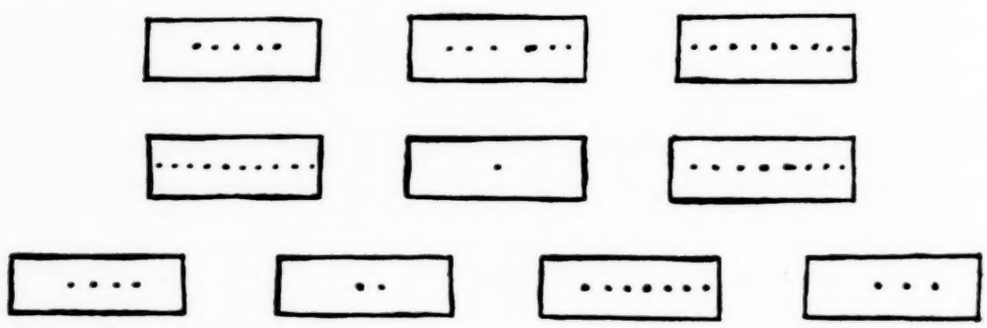
The cards were shown in the following predetermined, though not chance, order:

4, 7, 1, 9, 6, 2, 10, 3, 8, 5.

The controlling criteria in such an arrangement were: (1) to prevent reasoning out the number shown, as would be clearly possible if any arrangement such as that of ascending or descending order were used, or if there were present any other clues which would enable the child to intelligently guess in a manner not dependent upon his concept possession; (2) to provide considerable intervals between successive numbers, say 4, 5, and 6, or 7, 8, and 9, etc., in order to reduce the possibility of judgment by direct comparison and to make it more possible for a youngster to give the same estimate for two or more different stimuli by providing intervals for allowing their responses to retire from consciousness between numbers to which he would likely give the same estimate.

No announcement was made to the child as to the range of the numbers of dots to be shown him. No suggestion was given him that he might not be shown the same number of dots twice. Effort was made to relieve the child of timidity and nervousness at the outset and while he was given no definite clue as to the success or failure of his estimates, he was made to feel always that he was 'getting along all right.'

The second test is a variation of the first. We shall refer to it as the 'dot selection' test. The same material was used, the cards containing the dots. The cards were now arranged on a table as shown in the diagram below:



The pupil was asked to designate the card bearing the following numbers of dots in the following order:

2, 9, 5, 1, 10, 3, 6, 8, 4, 7.

Here again the child was not permitted to count. If the child seemed to begin to count his attention was distracted and he was asked not to count.

The third test we will call the 'marble' test. This test was given as was the third giving of the 'dot recognition' test, a month after the other tests. Only 26 children were given these two tests as but that number of the original 40 were available at the time these tests were given. In this test the experimenter displayed instead of cards containing dots, marbles in her hand grouped by chance, that is as they happened to group themselves when the hand was opened. The numbers of marbles shown were followed in order as in the 'dot recognition' test. The ages of the youngsters at the time of the final testing was obtained and recorded from the school records.

In scoring such tests there are several possible procedures. The simplest procedure would be to credit a child with possessing a concept of a given number if, and only if, he responded correctly to that number stimulus. There are certain uses to which the data may be put for which such a simple scheme of scoring might be ample. However, it is desirable in the consideration of many questions arising in a study of this nature to have some means of measure of any appreciable concept-possession and not to be limited to measurements which consider only 100 per cent. accurate concepts. It is quite evident that a youngster who estimates 8 for 9 has a fairly adequate concept of 9, one might say almost as good a concept as one who responds correctly. Many intelligent adults might estimate 8 for 9 if the estimations were not done under very favorable conditions. So it seemed very desirable to devise a means of scoring which would give adequate value to all guesses which, while not exactly accurate, indicated some degree of possession of number concept. It would seem at first that some proportional scheme would be usable. Missing a large number,

say 9 or 10, by 1 apparently indicates a higher degree of concept possession than missing a smaller number, say 1 or 2, by 1. Perhaps it might be fair to say that a guess of 4 for 5 is approximately equivalent for our purposes to a guess of 8 for 10. There are several difficulties, however, connected with the use of such a plan. In the first place, the location of a degree of deviation, beyond which all guesses indicate no concept at all, does not seem easy to fix, and again it is quite difficult to justify any proportional scheme mathematically. It is quite possible that the relation of miss to stimulus is not linear. We have no ready means of determining whether it is or is not linear, or, if it is not linear, just what the function is.

In order to devise a serviceable means of evaluating estimates it was necessary to make several rather justifiable assumptions. It would seem quite sensible to assume that a guess that was no nearer to the correct number than a chance guess would indicate a lack of number concept number more nearly than a chance guess is symptomatic of number concept. We would need to know how near such children would guess by chance to the respective numbers. If one could ascertain the respective deviations of a fairly large number of chance guesses of these children to, let us say, the number 4, one could calculate the median of these deviations and chance guesses would be as likely to be within as beyond the limits determined by that constant. One would be entitled to say that guesses of a deviation less than the median deviation were symptomatic of number concept; that for a fairly large number of guesses chance elements would tend to compensate each other.

With these considerations in view, there were obtained from these youngsters 238 guesses at the number of dots that might be on the back of a blank card. It will be obvious to the reader that the same set of guesses may be used to determine the median deviation of chance guesses from any number. As the children had no idea of the number of dots on the back of the card, (indeed, there were none), any one number may be assumed to be the number guessed at, as well as any other.

Below is a tabular distribution of the guesses:

TABLE I

DISTRIBUTION OF GUESSES AT UNKNOWN NUMBER

No. guessed	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Frequency	10	15	13	20	21	23	15	17	16	26	9	9	7	3	3	7
Percentages	4	6	5	9	9	10	6	7	7	11	4	4	3	1	1	3

No. guessed	17	18	19	30	31	32	33	34	35	40	69	90	100	Total
Frequency	3	4	5	1	2	1	0	1	1	1	1	1	3	238
Percentages	1	2	2	0	1								1	

The median deviations of these guesses from the numbers 1 to 10 in order are (to the nearest tenth):

No.	1	2	3	4	5	6	7	8	9	10
M.D.	6.6	5.6	4.6	3.6	3.3	3.3	3.2	3.2	3.9	4.3

One may say that a guess at 4 deviating from 4 by 3.6 or more is indicative of no concept possession as regards 4. A child of the age under consideration should in 50 per cent. of the cases do that well by sheer chance guess. A guess at '4' deviating by less than 3.6 is closer than one would expect the children of this level of development to come by chance.[1] Inspection of the median deviation of the chance guesses from the other numbers brings out a curious fact. It will be noted that the median deviations of guesses from 3 through 10 vary only from 3.1 to 4.6 with an average of 3.85, but that those of the guesses from 1 and 2 are significantly higher. A consequent of this fact, when viewed in the light of our

[1] It should be apparent that, while the error involved in assigning value to any one guess on this basis is quite large, too large for use in evaluating single guesses on this basis this error diminishes very rapidly with the increase in the number of guesses, until where $N = 40$, as in this study, it is practically negligible, *i.e.*, $\text{P.E.}_m = \frac{.6745\sigma}{\sqrt{N}}$ $= \frac{.6745 \times 1.9}{\sqrt{40}} = .2-$ or in terms of the graduation of from 0 to 1, $\frac{.2}{3.85} = .05$.

assumptions with respect to the value of an estimate in terms of chance, is somewhat startling, namely, that we are led to conclude that any given miss in estimating 1 or 2 is decidedly more symptomatic of concept than a miss of the equal deviation in estimating any of the larger number. For example, it would follow that a guess of 3 for 1 should be scored much higher than a guess of 12 for 10. This seems difficult to justify. Why this is true may be seen however from the following graph of the distribution of guesses.

CHART I. Frequencies of appearance of numbers in blind guesses

We know that the median deviation taken from modal and closely adjacent regions in distributions resembling that of error or normal distribution are certain to be less than such taken from points of lying in areas of smaller frequency or of lesser proximity to modal regions. Examination of the above graph would indicate smaller median deviations from 3 through 10 and larger ones below this area and very much larger ones as the point of deviation moves up along the upper tail of the graph line. Children are more or less limited, and are more familiar with numbers of 1 to 10. Their guesses, as a consequence, bulk largely within these limits.

At first consideration the idea of scoring a guess of 2 for 1 higher than one of 10 for 9 would seem ridiculous and yet upon careful consideration much may be said for doing just that thing. In the first place, as an inspection of Table I. will show, the youngster has 59/25 times as many chances to guess 9 with a miss of not more than 1 as he has to miss 1 by not more than the same deviation. By mere chance [1] he is much

[1] By chance is meant those errors independent of number concept. Chance guesses may be influenced by favorite numbers or most familiar numbers and other such factors.

more likely to guess closely to numbers 6, 7, 8, 9, and 10, than to 1 or 2. In this case he is 2.36 times as likely to miss 9 by not more than 1 as to miss 1 by not more than 1. That is the same as saying that if a youngster guesses 3 or 4 for 2, or 2 or 3 for 1, it seems much better evidence that he has some concept of 2 or 1, that would be a guess of 6 or 7 for 9, since one's guess is more likely by chance to range along about 6 or 7 than in the neighborhood of 2, 3 or 4.

If there had been numbers extending below 1 for some distance it is quite likely that the frequency of guesses of smaller size would have been much larger and median deviations in this region much smaller. It is quite possible therefore that the zero points of 1 and 2, namely, the points of median deviation from these numbers which are identical in this particular instance, 7.6, should be arbitrarily drawn in. In fact it will seem to many that still better would be some scheme of weighting guesses proportionally, that is a scheme by which larger misses in estimates of larger numbers would correspond in value to smaller misses in estimating smaller numbers. Were we interested in measuring what could be expected of people in estimating, *i.e.* relative merit, such a scheme would be advisable but our problem is concerned with absolute rather than relative performances. Fortunately in this study the question of weighting misses in estimating 1 or 2 is of academic importance only, since for our subjects there were no misses on these numbers. For the remaining numbers the median deviations are so nearly uniform that error in scores consequent to assuming uniformity at some average is quite small. (See tabulation on page 451.)

The average value of the median deviations from numbers 3 through 10 having the least maximum deviation from the obtained median deviation is equal to the average of the maximum and minimum median deviations $\frac{4.6 \text{ plus } 3.1}{2}$ which is 3.85. So, if an estimate deviates from the number being estimated by 3.85 or more we are justified in saying that the estimation has no appreciable concept of the correct number. We will assign a value of 1.00 to a correct estimate, a value of

0 to those inaccurate by 3.85 or more. Intermediate approximates range between 0 and 1. An assumption of a linear scale between these limits is as reasonable as any other and has the added merit of having the least error from other possible distributions of value or merit which might be assumed. Perhaps if any deviation from linearity is to be made, it should be in the direction of arbitrarily weighting slightly close estimates containing small error and underweighting estimates containing large error.

A linear scale with zero points at 3.85 would give us the following equivalents or weights.

Miss of	0	1	2	3	4 and over
Weight of	1	.741	.481	.221	0

The weights actually used in scoring the estimates were respectively 1.0, .8, .5, .2, 0.

3. Number Concepts Possessed by Kindergarten Children of Ages 4½ to 6 Years

From the data it seems clear that normal kindergarten pupils of ages 4 years 6 months and more all know and recognize 1 and 2 to an extent approximating perfection. It is the exceptional child that does not have the same adequate concept of 3—from our data one child in seven (6 of 40)—and of these but 1 in 20 is not able to know three as such, when presented as frequently as three times in five.

Table II

Percentage of Kindergarten Pupils Estimating Population of from 1 Through 10

	1	2	3	4	5	6	7	8	9	10
Five successful attempts	100	100	85	15	0	0	0	0	0	0
Three successes in five attempts	100	100	95	65	12	0	0	0	8	12

In Table II. are given the percentages of these youngsters who were able to estimate accurately the frequency of groups

or rather lines of dots from 1 to 10. The first row of figures is based on a standard of five successes in five attempts, the second on a standard of three successes in five attempts.

When it is considered that a portion of the correct estimates were not estimates at all, but lucky guesses, it seems quite easy to conclude that kindergarten youngsters with occasional exceptions do not have *accurate* concepts of numbers above 4 and that their concepts of numbers above 4 are so indefinite that rarely can one correctly estimate numbers above four as often as three times in five.

One should not however construe these results to mean that youngsters of this age have *no* concept of numbers above 4. Quite the contrary is true. Many youngsters have sufficiently adequate concepts of these larger numbers to enable them to estimate them invariably with small error and there is a large number of them who can estimate them within one a majority of times. These facts are clear-cut evidences of appreciable development of concept. This data and further discussion of this point is given further on.

We will now turn to the consideration of data obtained from single trials of the three different tests. The frequencies of correct estimates is given in Table III.

TABLE III

NUMBER OF RIGHT ESTIMATES ON THE THREE TYPES OF TESTS

	1	2	3	4	5	6	7	8	9	10
Dot-recognition (40)	40	40	34	28	10	10	7	7	6	5
Dot-selection (40)	40	40	31	19	14	9	8	10	5	10
Marble test (26)	26	26	31	13	19	13	13	16	16	14

The numbers in parentheses refer to the numbers of pupils tested. However, not all of the correct 'estimates' may be attributed to possession of concept. There were no doubt many accidental hits. For example, upon reference to the distribution of chance guesses given in Table I., we find that 11 per cent. of chance guesses were 10. In the first giving of the 'dot-recognition' test in estimating 10, there were 5 rights among 40 children. If all estimates had been

sheer guesswork 4.4 per cent.[1] of these on the average would have been correct. However we have no way of estimating how many of these estimates were sheer guesswork. By a variation of the method used in scoring multiple-choice tests we could say that if one guess in nine would be correct by chance the number of real accurate estimates, after allowing for chance hits, would be given by the following formula:

$$\text{Score} = R - \frac{W}{8}.$$

In the instance referred to above we would obtain

$$5 - 35/8 = 5/8$$

correct estimates not due to chance. We would be justified in concluding that not more than one child in 40 or approximately 2 per cent. of children of the age studied could be expected to recognize 10 dots under the conditions of an experiment. Similar calculations for each number and for the first giving of the 'dot recognition' test, for the 'dot selection' test, and for the marble test are shown in percentages in Table IV.

TABLE IV

PERCENTAGES OF CHILDREN WHO MIGHT BE EXPECTED TO HAVE CONCEPT OF NUMBERS 1 THROUGH 10 AS TESTED BY THE 'DOT-RECOGNITION' TEST, 'DOT-SELECTION' TEST AND THE MARBLE TEST

	1	2	3	4	5	6	7	8	9	10
Dot-recognition	100	100	83	68	18	17.5	12	11	9	2
Dot-selection	100	100	76	42	28	17	13	19	6	16
Marble test	100	100	82	45	29	12	16	18	18	05
Average	100	100	84	55	25	12	10	16	11	8

Here again we have evidence that the number concepts of kindergarten children extend uniformly through two, approximately so through three, and while about as frequently as not the child 'knows' four, it is the exceptional child who has an accurate concept of numbers above four. Yet it is evident from this representation that children of this age do

[1] See Table I.; 11 per cent. of 40 = 4.4.

possess some degree of concept of each of the numbers even though it is not sufficiently accurate to enable them to estimate accurately the population of groups.

Light is thrown on the question of how well children actually know numbers above three by a comparison with the mean deviation of the misses of the respective numbers. These are given in Table V.

TABLE V

DISTRIBUTION OF DEVIATIONS OF ESTIMATES FROM CORRECT NUMBERS (3 TO 10)

No.	No. of Rights	Size of Deviation from Correct Population										Mean Dev.	Chance Median Dev.
		1	2	3	4	5	6	7	8	9	10 Median		
3	92	6	5	2				1			.14	.3	4.6
4	60	26	10	2	2	2	2	1	0	1	.4	.9	3.6
5	34	39	6	12	8	6	0	0	2		1.0	1.6	3.3
6	22	43	18	16	5	1	1	0	0	0	0 1.2	1.6	3.3
7	16	46	20	20	1	1	0	2	0	0	0 1.3	1.7	3.1
8	23	34	26	10	7	3	0	0	2	0	1 1.4	1.8	3.2
9	17	42	24	11	2	4	5	0	0	0	1 1.4	2.1	3.9
10	19	26	26	14	8	8	2	1	0	0	2 1.8	2.25	4.3

A comparison of Table V. with Table II. will seem to demonstrate the inadequacy of judging the possession of number concepts by data which recognizes only accurate estimate. As will be seen from a comparison of the last column of Table V. with the two immediately preceding columns, children on the average estimate much more closely to numbers one to ten than would result from mere chance. It must be noted that again there is evidence that children of this age have much better concepts of three and four than of numbers beyond four. Table VI. also gives weight to these approximate estimates. For each number from 1 to 10 the mean 'miss' of each child has been computed from his estimates in the three tests, the first two givings of the dot-recognition test and the dot-selection test. This miss has been translated into scale values with zero points at misses of 3.85 or larger as previously described. The mean of the individual scores so obtained for each number from 1 to 10 is given.

TABLE VI

THE MEAN SCORE OF 40 KINDERGARTEN CHILDREN AS DERIVED FROM THE MEAN ERRORS OF ESTIMATE

Number	1	2	3	4	5	6	7	8	9	10
Av. score	1	1	.89	.79	.67	.65	.61	.56	.58	.49

There are no directly comparable results from other investigations. Differences in method of experiment and the nature of the data in ages of children, selection, etc. make comparisons impractical.

A study of concepts of children entering school conducted in 1869 under the direction of the Berlin pedagogical Verein included 2,238 children. In this study the results obtained indicated that but 74.35 per cent. of such children knew two, but 74 per cent. knew three and but 72.65 per cent. knew four. Superintendent Seaver in 1880 under the direction of G. Stanley Hall found of the children in the kindergarten classes in the Boston schools that approximately 93 per cent. knew 'three,' 88 per cent. 'four,' and 21 per cent. 'five.' The differences in the possible conclusions from these two studies is significant. In the German study, 74 per cent. and 72.65 per cent. knew three and four respectively. In the Seaver study these percentages are 93 per cent. and 88 per cent. respectively.[1]

Lietzmann (27) states that when the child comes to school he will in a majority of cases already know number words up to 10 or 12. Of 1,217 children entering Volkschule in Breslau, 90 per cent. could count to 5, 78 per cent. could count to 10, 4 per cent. to 100, and about 1 per cent. over 100.

4. THE DEVELOPMENT OF NUMBER CONCEPT WITH AGE

The number of children is so small that when separated into age groups the results are subject to sharp influence of individual fluctuation, yet such results are indicative and distinctly superior to the 'single child' studies which constitute the only data we have had on the problem. For con-

[1] These studies are reported in brief by Hall in "Aspects of Child Life and Education," Ch. I. (22).

venience the children are divided into four groups with age intervals of six months each. Children are classified in these age groups so that the class index represents the midpoint of the range of the class group, *e.g.*, a child of 4 years 10 months is classed as a 5 year old. The mean scores and mean number of right guesses for each age is given in Table VII. (See Chart II.)

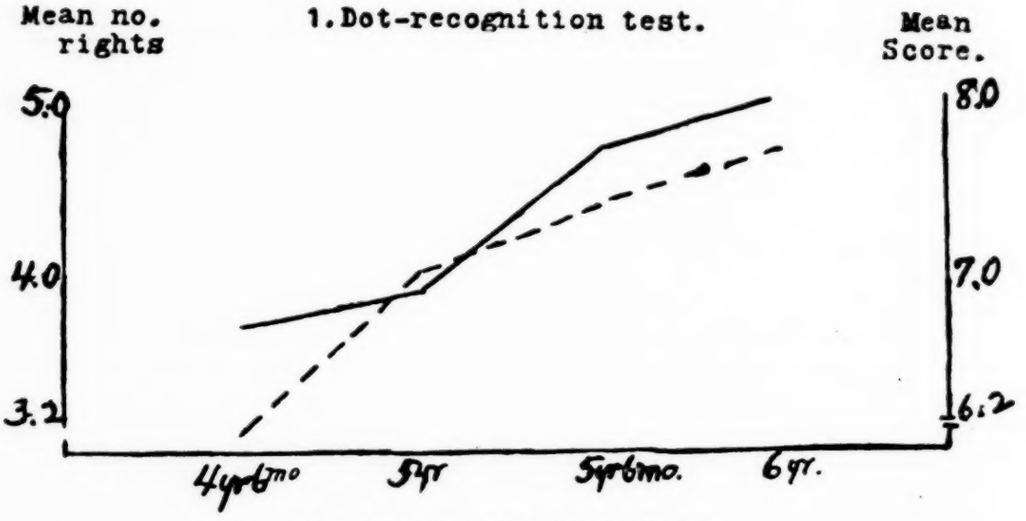
TABLE VII

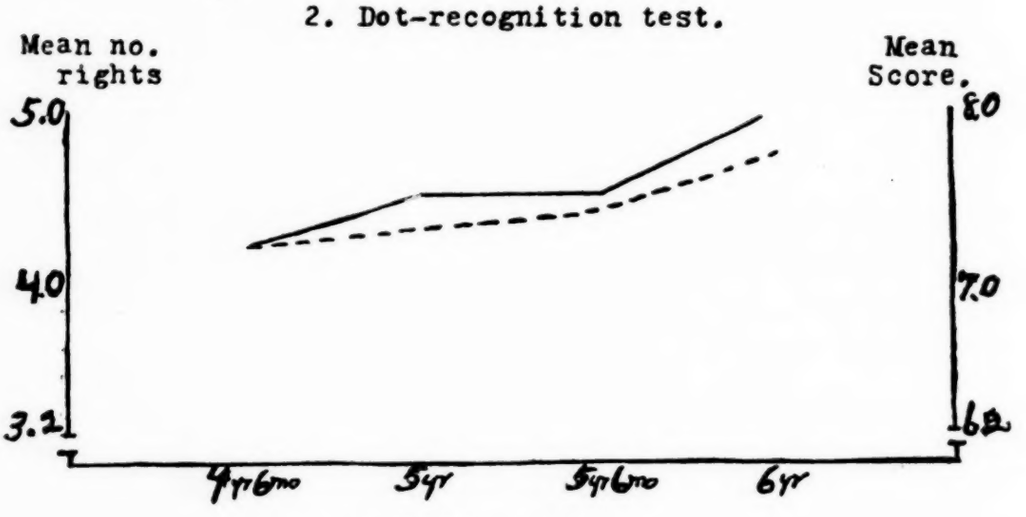
NUMBER PERFORMANCE OF FOUR AGE LEVELS OF KINDERGARTEN CHILDREN

Age	Dot-Recognition			Dot-Selection			Marbles		
	No.	Av. Right	Av. S.	No.	Av. Right	Av. S.	No.	Av. Right	Av. S.
4 yr., 6 mo.	4	3.7	6.1	4	4.2	7.2	3	3.7	7.4
5 yr.	11	3.9	7	11	4.5	7.3	5	4.8	7.8
5 yr., 6 mo.	13	4.7	7.4	13	4.5	7.4	10	4.7	7.97
6 yr.	12	5	7.7	12	5	7.7	8	4.6	7.3

An examination of Table VII. discloses that apparently the development of number concept as would be expected is a function of chronological age and its concomitant environing influences. In the dot-recognition test each successive level of 6 months age intervals is represented with a higher score than the previous. That this is not strictly true in the data from the dot-selection and marble tests is no doubt in part due to individual fluctuation which in groups of 9 to 13 children, as we deal with here, is likely to influence averages unduly and in part to a selection at the older age levels which is effected by the promotion of the brighter older children into the first year. The lack of superiority of the 5 year 6 months old group over the 5 year olds and the actual inferiority of the 6 year olds to the 5 year 6 months group in the marble test will yield to satisfactory explanation only after detailed study of the individuals constituting this group.

A composite grouping of the results of all three tests by means of taking as values in the derived table the arithmetical means of the comparable entries in the three rows of Table IV. is given in Table VIII. The error resulting from this arbitrary procedure is very likely to be much less significant than the





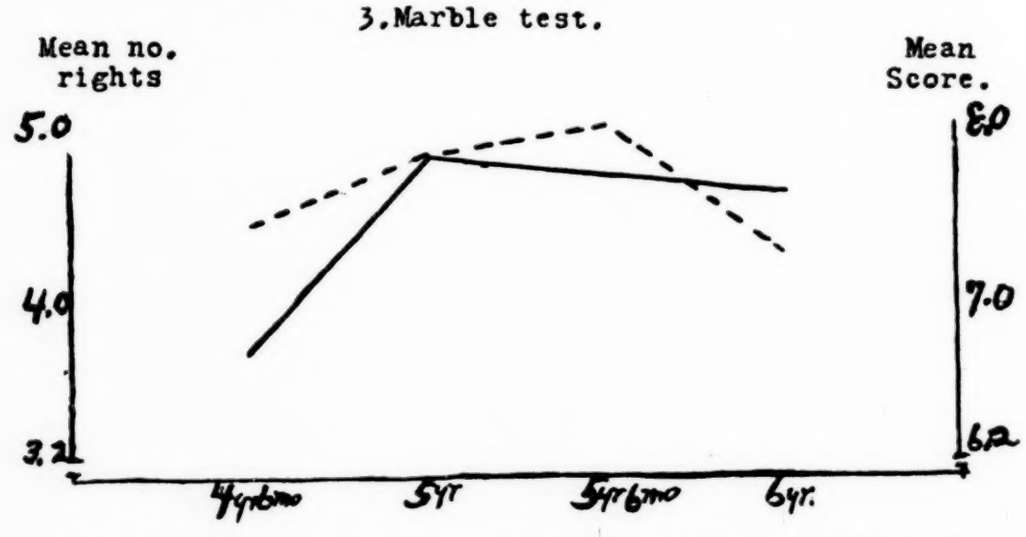


CHART II. Number performance of kindergarten children of four age levels

error of sampling due to the limited number of children at each age group.

TABLE VIII

AVERAGE SCORE IN RIGHTS AND WEIGHTED SCORE FOR FOUR AGE LEVELS OF KINDERGARTEN CHILDREN

Age	Pop.	Rights	Scores
4 yr., 6 mo.	4	3.9	6.9
5 yr.	11	4.4	7.4
5 yr., 6 mo.	13	4.6	7.6
6 yr.	12	4.9	7.6

The presence of a selective factor in the older children tends to obstruct complete demonstration; the limited range in age in the group studied operates to restrict the study of this growth curve to the small age range of approximately 20 months. To say whether children of a certain age know various numbers cannot be done without a certain amount of arbitrary procedure. It is clear that children of 4 years and 6 months and older know 1 and 2. Beyond this one can only say that children know the numbers from 3 to 10 with varying degrees of accuracy, varying with the child, varying with the number. If we set as an arbitrary standard an average miss of not more than 1 which corresponds to scale score of .8 we find the following average performance for the age groups: (See Chart III.).

TABLE IX

PERCENTAGES OF CHILDREN KNOWING NUMBERS 3 TO 10 BY AGE GROUPS
(Standard = average miss of not more than 1 on three attempts—
2 of dot-recognition—1 of dot-selection)

Age	N.	Percentage Averaging Misses not Greater than 1							
		3	4	5	6	7	8	9	10
4 yrs., 6 mo.	4	75	50	50	50	50	50	25	0
5 yrs.	11	91	54	55	18	00	36	18	9
5 yrs., 6 mo.	13	85	70	53	45	30	47	38	30
6 yrs.	12	83	83	67	50	42	25	42	33

Because of the smallness of the proportion of the 4 year 6 months group the effect of which will be noted in the entries

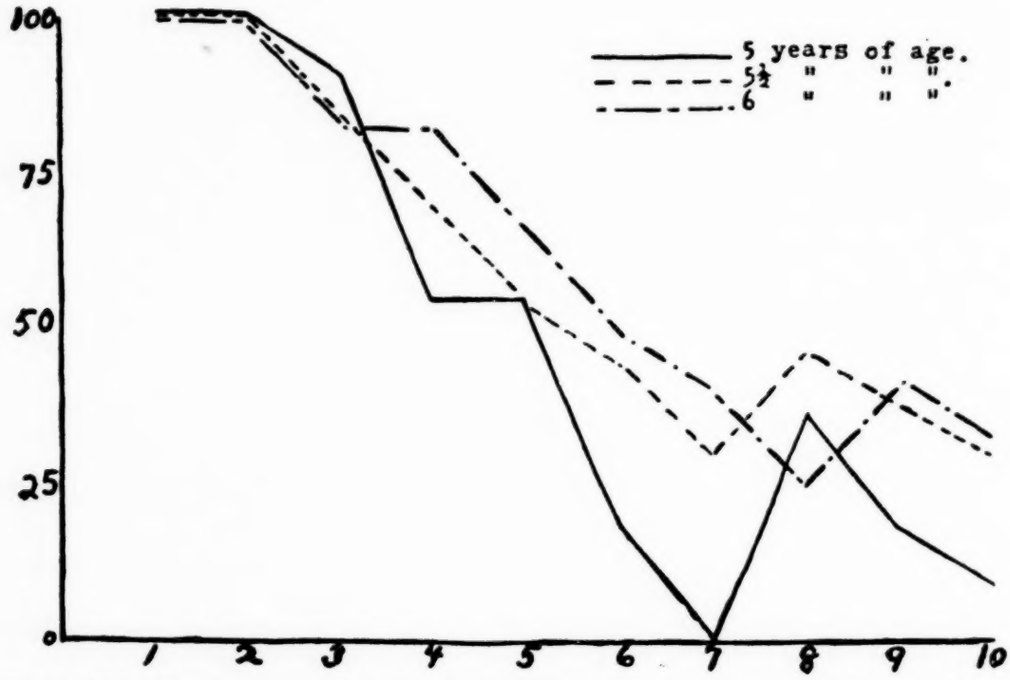


CHART III. Percentages estimating population of dot-groups at estimating age levels. Standard: Two successes in three attempts

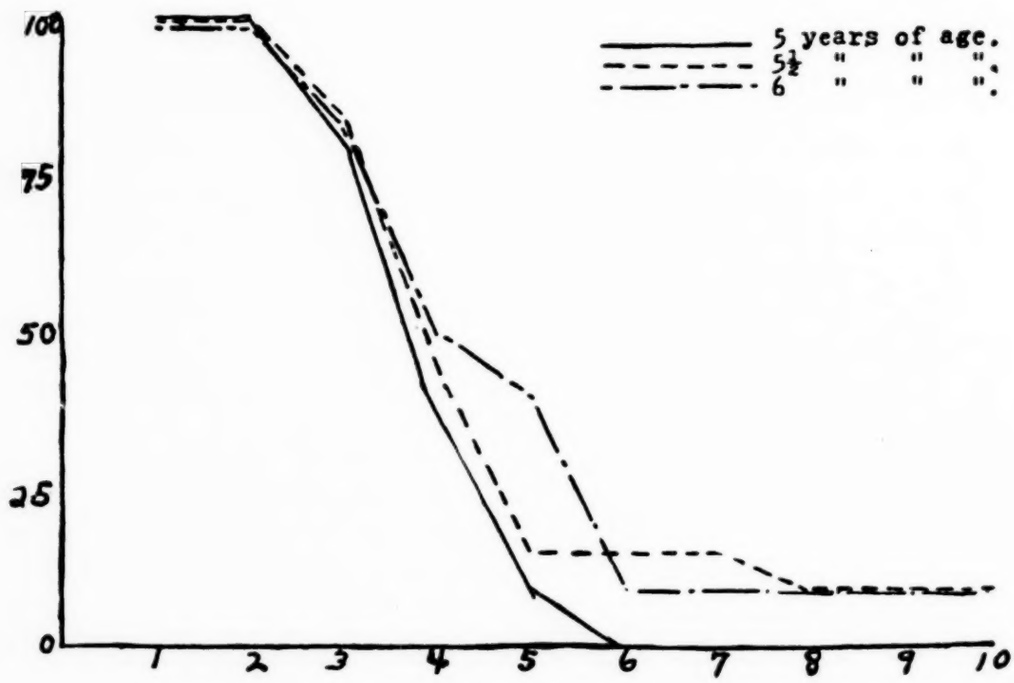


CHART IV. Same as above, standard: average miss of not more than 1.0 in three estimates

for that group practically no significance can be attached to those entries.

Similar tables might be constructed for various arbitrarily determined standards but in each such tabulation there is neglected the evidence of concept of furnished by less accurate performance which is however more accurate than chance guesses and in each no weighting is given the more accurate estimates as against the less accurate of those exceeding the particular standard set.

If we set as a standard two accurate estimates for three attempts we obtain the percentages given in Table X below.

TABLE X

PERCENTAGE OF CHILDREN ESTIMATING NUMBER OF RESPECTIVE DOT-GROUPS AT SUCCESSIVE AGE LEVELS

(Standard = two successful estimates in three attempts)

Age	Pop.	1	2	3	4	5	6	7	8	9	10
4 yr., 6 mo.	4	100	100	75	25	0	0	25	0	0	0
5 yr.	11	100	100	81	36	9	0	0	0	0	0
5 yr., 6 mo.	13	100	100	85	45	15	15	15	8	8	8
6 yr.	12	100	100	83	50	42	8	8	8	8	8

Children of the fifth year possess very accurate concepts of 1, 2, and 3 and rapidly forming concepts of 4. During the sixth year the concept becomes sufficiently well-formed that the median 6-yr.-old child will respond accurately two times out of three or better to four, and almost as well to five.

5. SEX DIFFERENCES

No significant sex differences were found. The number of pupils tested and the range of age were too small to permit of any reliable comparison of growth between the sexes. Table XI. furnishes a comparison of the number of rights and the weighted scores of the two sexes as determined from each of the three tests.

The superiority of the girls is uniform though in general of an insignificant amount. Some attention however should be given to the superiority in the dot-recognition test. Here the difference if judged by rights is equivalent to a six month's

development as shown in Table III. and if judged by weighted scores a three month's development. How much of this may be explained by the willingness of the girls to try harder is difficult to say, as we have no objective evidence that they did try any more though the experimentor is inclined to think that they did. It would seem that if such were the explanation for the sex difference in the dot-recognition test results that there would be like differences in the other tests. As the mean and range of ages of the two sexes are approximately equal, the girls younger by .5 months than the boys, the difference can hardly be attributed to difference in maturity.

Table XI

Mean Number of Correct Estimates and Weight Scores of Boys and Girls for Three Tests Separately

	Dot-Recognition		Dot-Selection		Marbles	
	Rights	Scores	Rights	Scores	Rights	Scores
Boys (18)	4.2	7.2	4.6	7.2	4.3	7.3
Girls (22)	4.7	7.5	4.7	7.7	4.5	7.4

Note: The performances of 13 boys and 13 girls only are included in the marble test data.

If we compare the sexes on the basis of the second and third giving of the dot-recognition tests, the sex difference is not clear-cut. It seems quite likely that the apparent sex difference very largely disappears in the chance fluctuations of individual performance.

Table XII

Mean Number of Correct Estimates and Weighted Scores by Sexes, Second and Third Giving of Dot-Recognition Test

	2d Giving		3d Giving	
	Rights	Scores	Rights	Scores
Boys	4.7	7.3	4.5	7.3
Girls	4.7	7.7	4.7	7.4

6. Reliability and Validity of the Tests Used

There is no criterion by which we can judge the validity of these tests. The most trustworthy information we have regarding the number concepts of these children tested is that gotten through the use of the tests themselves. The question of their validity must be left to logic. Some light however is thrown upon the question of the ability of the tests to measure the possession of number concepts of 1 through 10 in general as opposed to the mere ability to measure such in a special situation. Freeman and others have shown that variation of the grouping of objects in perception experiments tends to produce variation in the results obtained. One may have demonstrated a concept of five when five objects are grouped in one fashion and not when grouped differently. Or we may say that one may recognize five dots but not five pigs and may not be able to assemble a group of five objects or to select a group of five objects from other groups of other frequencies. This problem in itself constitutes a problem for extended and careful study and can be considered here only as it may afford data on the question of the validity of the tests. Some idea may be obtained of whether any one of the tests measure the same functions as the others. If we find that such is the case for any one test, then as far as the three tests sample the general field of number concept, that far can we assume the given test to measure it.

As a measure of reliability the coefficient of correlation between two forms is superior to that constant calculated from two givings of the same test. Certain of the objections to the use of the coefficient of correlation between two givings of the same test, however, do not apply with much force to this particular test.[1]

The intercorrelations from which inferences may be made as indicated in the preceding paragraphs are given in Tables XV.

These intercorrelations indicate a sufficient reliability of the tests for reasonably safe use of data obtained with them.

[1] Memory carry over, learning from the test itself, etc.

The slightly smaller correlation of the third with the second giving of the dot-recognition test as compared with that obtained between the first and the second givings may be accounted for by the fact that the first two were given the same day and the third a month later during which time it is reasonable to assume that learning has taken place unevenly among individuals.

TABLE XV

INTERCORRELATIONS BETWEEN WEIGHTED SCORES ON THREE GIVINGS OF THE DOT-RECOGNITION TEST AND ONE EACH OF THE DOT-SELECTION AND MARBLE TESTS

	Dot-Recognition (1)	Dot-Recognition (2)	Dot-Selection
Dot-Recognition (1)		.653±.091	.510±.171
Dot-Recognition (3)	.584±.104		
Marbles	.404±.164		.357±.171

The relatively high intercorrelations between the different tests—.51, .404, and .357, when compared with the reliability coefficients of .584 and .653 indicates that the tests are testing the same general or very concomitantly developing functions. The extent to which the tests tap the same or similar functions is, as would be expected, less when we consider the dot-recognition and marble tests than when we consider dot-recognition and dot-selection owing to the similarity of the latter pair of tests.

7. SUMMARY AND CONCLUSIONS

There is comparatively little known regarding the number concept with children of pre-school age. The chief source of available data is a number of intensive studies of individual children, chiefly vocabulary studies. The results of these are of little value because of their incompatability, the small number of cases, the lack of a check on possible selective factors and because the data resulting are difficult to interpret with respect to the problem of number concept. Practically all other number studies of children are confined to children who are able to write and to take group tests.

One reasonable way to test for number concepts is to

test for the ability to estimate the number of objects in a group. There is no end to the possible variation of such tests. As reasonable as any are tests which measure the ability to estimate the number of plainly visible dots on a card exposed just long enough to permit clear sense image but not long enough to permit counting, which measure the ability to select a card bearing the requested number of dots from an assortment of cards bearing various numbers of dots, and which require the estimating of the number of objects, say, marbles in a group of such.

In the scoring of such tests, a device should be used which measures concepts of whatever accuracy from concepts which are not sufficiently formed to enable the possessor to make better estimates than he would be able to make by chance, if he had no concept at all, up to concepts which enable him to make accurate estimates. Such a device approximately was the plan of equating estimates on linear scale so that perfect estimates were given a value of 1.00 and estimates deviating from true estimates by 3.85 (mean deviation of chance guesses from the number guessed at) were assigned a value of 0 and misses of less than 3.85 a proportional value between 0 and 1.0. The results of such a test given to 26 to 40 kindergarten children of ages from 4 years 6 months to 6 years give grounds for the following conclusions: 1. Children of this age range have completely accurate concepts of 1 and 2, very serviceable and accurate concepts of 3 and a very serviceable concept of 4, and of 5, 6, 7, 8, 9, and 10 rather vague concepts though serviceable to a slight degree (children on the average being able to estimate within from 1.2 to 2.1 of these numbers). (See Tables II., IV. and V.)

2. The development of number concept is a function of age and the concomitant educational development of children. (See Tables VII., VIII., IX., and X.)

3. Very slight sex differences, corresponding at most to probably not more than 3 months development, favor the girls. Sex difference in any significant degree appears only in the dot-recognition test.

5. The dot-recognition test has a reliability of approximately .60 to .65 for narrow ranges,

(r_{11} = .653 on tests given twice one day,
r_{11} = .584, tests given with interval of a month.)

6. The intercorrelations furnish strong evidence that the three tests are tapping the same or similar functions and that recognizing number groups is not essentially different from estimating the population of groups, at least as far as numbers 1 to 10 and kindergarten youngsters are concerned.

Bibliography

1. Alling—Aber, Mary R. An Experiment in Education. New York: Harpers, 1897. 24 pp.
2. Arnett, J. D. Counting and Adding. Minor Studies from Psych. Lab. of Clark University—*Amer. Jour. of Psych.*, *16*: 327–336, 1905.
3. Badanes, Saul. The Falsity of the Grube Method of Teaching. N. Y. University thesis, 1895. 47 pp.
4. Ballard, P. B. The Teaching of Mathematics in the London Public Elementary Schools. Coeduc. Dept. Great Britain Special Reports, Vol. 26.
5. Beyer, Thos. P. Vocabulary at Two Years. *Ed. Review*, *49*: 401–413, 1912.
6. Boas, Franz. Language and Thought. Intro. to Handbook of Amer. Institute of Lang. Washington; Gov. Phy. Office, 1911. 100 pp.
7. Boyd, Wm. Development of a Child's Vocabulary. *Ped. Sem.*, *21*: 95–124, 1914.
8. Brandenburg, George, C. The Language of a Three-Year-Old. *Ped. Sem.*, *22*: 89–104, 1915.
9. Burnett, C. J. The Estimation of Number. Harvard. Psychological Studies, II., 1906, 349–404.
10. Conant, Levi, L. The Number Concept. New York: The Macmillan Co., 1896. 218 pp.
11. Bush, Arthur O. Vocabulary of a Three-Year-Old Girl. *Ped. Sem.*, *21*: 125–140, 1914.
12. Court, Sophie, R. R. Numbers, Time and Space in the First Five Years of a Child's Life. *Ped. Sem.*, *27*: 71–84, 1920.
13. Dearborn, G. U. Motor-Sensory Development. Baltimore; Warwick and York, 1910. 215 pp.
14. De Croly, Dr. M. Le, and Degand, Mlle. Julia. Le Evolution des notions de quantity chez l'enfant. *Archives d. Psych.*, Vol. 12, 1912, 88–121.
15. Dewey, Jno. Some Remarks of the Psychology of Number. *Ped. Sem.*, *5*. 426–434, 1898.
16. Freeman, Frank, N. Grouped Objects as a Concrete Basis for the Number Idea. *Elem. School Teacher*, *12*: 306–314, 1912.
17. Freeman, Frank, N. Psychology of the Common Branches. Houghton Mifflin Co., 1916. 271 pp.
18. Freeman, F. N. Untersuchungen über den Aufurerksamkersamfang und die Zahlauffassing. *Pedagogische-Psychologische Arbeiten*, *1*: 88–168.
19. Grant, James R. A Child's Vocabulary and Growth. *Ped. Sem.*, *22*: 183–203, 1915.

20. Groszmann, Maximilian P. E. The Career of the Child. Boston: R. E. Badger, 1911, 335 pp.
21. Guillet, Cephas. The Growth of the Child's Concepts. *Ped. Sem.*, *24*: 81–96, 1917.
22. Hall, G. Stanley. The Contents of Children's Minds on Entering School. Chapter I. in Aspects of Child Life and Education. New York: D. Appleton & Co., 1908.
23. Hall, Mrs. W. S. The First Five Hundred Days of a Child's Life. *Child Study Monthly*, *2*: 586–608, 1897.
24. Howell, Henry Budd. A Foundational Study in the Pedagogy of Arithmetic. New York: The Macmillan Company, 1914. 328 pp.
25. Langenbach, Mildred. The Study of a Five-Year-Old Child. *Ped. Sem.*, *22*: 65–88, 1915.
26. Lay, W. R. Fuhrer durch den duversten Rechenunterricht. Weisbaden, 1898.
27. Lietzmann, W. Stoff and Method des Richmenunterrichts in Deutschland. Leipsig: Quebners, 1912. 107 pp.
28. Leuba, H. The Personifying Passion in Youth. *The Monist*, *10*: 536–548, 1900.
29. Lucas, Mrs. A Child's Sense of Number. Child Study Monthly, 7: 5–7 and 8: 141–143, Feb. 1914 and Dec. 1915.
30. Lucas, Edward. Theorie of Number. Paris, Gouthes., 1891. 520 pp.
31. McLellan, J. A., and Dewey, Jno. Psychology of Number. New York: D. Appleton and Co., 1895. 395 pp.
32. McDougle, Ernest C. A Contribution to the Pedagogy of Arithmetic. *Ped. Sem.*, *21*: 161–194, 1914.
33. Major, David R. First Steps in Mental Growth. New York: The Macmillan Company, 1906. 360 pp.
34. Mateer, Florence. The Vocabulary of a Four-Year-Old Boy. *Ped. Sem.*, *15*: 63–72, 1908.
35. Messenger, J. F. Perception of Number. *Psych. Rev. Mon. Supp.*, V., 5, No. 5, June, 1903.
36. Meumann, Ernst. Vorlesunzen zur einfuhring in die experimental padogogich und die psychologischen grundlagen. Liepsig, Engleman, 1907. 2 vol.
37. Miller, Irving Edgar. Psychology of Thinking. New York: The Macmillan Company, 1909. 303 pp.
38. Moore, Kathleen Carter. The Mental Development of a Child. *Psych. Rev. Mon. Supplement*, No. 3, Oct., 1896.
39. Nanu, Helene A. Zur Psychologie die Zulil-Auffansusag. Wurzberg, 1904.
40. Nice, Margaret Morse. Development of a Child's Vocabulary in Relation to Environment. *Ped. Sem.*, *22*: 35–52, 1915.
41. Oatman, Miriam E. A Boy's Development at Eighteen Months. *Ped. Sem.*, *28*: 52–59.
42. O'Shea, M. V. The Psychology of Number—A Genetic View. *Psych. Rev.*, *8*: 371–383, 1901.
43. Perez, Bernard. The First Three Years of Childhood. London: Swan, Sonnensheim and Company, 1900.
44. Phillips, D. E. Number and Its Applications. *Ped. Sem.*, *5*: 221–281 (1897); also *5*: 290–298 (1898). Also Genesis of Number Forms. *Amer. Jour. of Psych.*, 8: 506–527, 1896.
45. Preyer, W. Mind of the Child. New York: D. Appleton and Company, 1889. 317 pp.

46. Ribot, Th. The Evolution of General Ideas. Translated by Frances A. Welby. Chicago: The Open Court Publishing Company, 1899. 231 pp.
47. Rowe, Eugene C., and Helen N. The Vocabulary of a Child at Four and at Six Years of Age. Ped. Sem., 20: 187–201, 1912.
48. Rusk, Robert. Experimental Education. London, Longmans, Green and Co., 1913. 364 pp.
49. Starch, Daniel. Educational Psychology. New York: The Macmillan Company, 1919. 473 pp.
50. Sully, James. Studies of Childhood. New York: D. Appleton and Company, 1896. 527 pp.
51. Tanner, Amy Elizabeth. The Child. Rand, McNally and Co., 1904, 1915. 532 pp.

INSTRUMENT FOR MARKING THE TEMPERATURE SPOTS ON THE SKIN

BY C. R. PENDLETON

University of Pennsylvania

While working on the skin in the Psychological Laboratories of the University of Pennsylvania, several difficulties were encountered in the endeavor to accurately localize the cold and warm points. The end bulbs of Krause, to which the cold spots are supposed to correspond, probably do not exceed one tenth of a millimeter in their largest dimensions, while the majority of them are smaller. The Ruffini cylinders sometimes measure as much as 1.35 mm. in length, but are elongated and fusiform.[1] No such minute localization is possible by the older methods, nor is such localization claimed for the method here described, but it is believed that this method is something of an improvement over the older methods.

The first difficulty is the size of the marks. Aniline dyes and fine camel's hair brushes make relatively large marks which are uneven in contour and tend to run. Marks made by pens are unsatisfactory for similar reasons. Solid points with ordinary fountain pen or draftsman's ink will not mark.

The second difficulty is that markers sometimes fail to work. Fountain pens are particularly bad for if they have stood for a few seconds they fail to write unless the point is large. Camel's hair brushes often dry up on the tip, and must be pushed down a bit further to make a mark, which results in a larger mark than desirable.

A third difficulty is the necessity to move the eyes from the point localized. If the marker fails to work, the eyes must be moved to correct the faulty condition. If the operator would make certain his marker will work he must dip the brush in the dye, or try the pen on paper just before marking.

[1] Piersol's 'Anatomy,' 1907, pp. 1016–17.

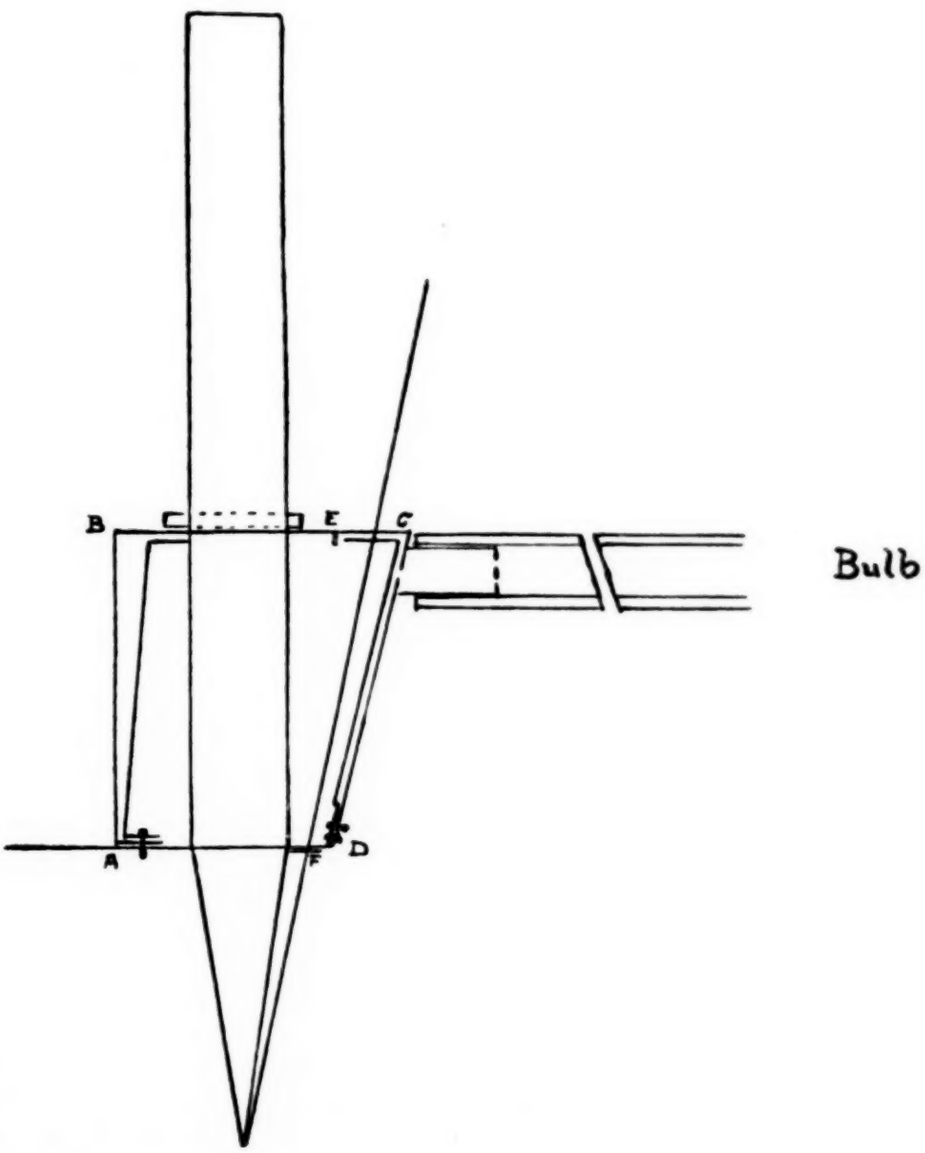
A fourth difficulty is the necessity to remove the dermic pencil and put the marker in its place. This involves error even when the skin is ruled with small squares, for even if the eyes see the exact point, it is often impossible to put the marking point accurately in place.

The beginning of a solution came with the use of printer's ink. If this be smeared thinly on a piece of glass, a blunt needle point can be 'inked' by touching the smear. It will not dry up nor fail to write for some time, possibly an hour; and the mark can be made as small as the eye can see, while there is no tendency to run.

Following the adoption of printer's ink and needle, an instrument was devised to carry the dermic pencil and the needle. On pressing a rubber bulb the needle drops, making a minute mark one half millimeter to the side of the warm or cold point. For purposes of calculation it is then a simple matter to make the necessary correction and thus have the subjectively indicated temperature spot accurately marked to within a degree of error probably at least not greater than one tenth of a millimeter. It has also been found that after all desired measurements have been taken the spots can be printed distinctly on paper, though in reverse order, by pressing a strip of paper against the skin.

The instrument is constructed as follows: (1) A frame *ABCD* is made by bending a strip of brass 2 cm. wide by 1 mm. thick into the form shown; one end is soldered at *A*, the other projecting to form a finger rest. Two holes are drilled in this, approximately 1 cm. in diameter, to carry the dermic pencil. (2) A strong spring, made of the same material, extends from *A* obliquely up to press on the pencil and hold it securely in place. (3) The dermic pencil is the usual variety. It must have a stop, a small brass bar which rests on the frame at *BC*, to hold it exactly at the right level. The point is one-half millimeter in diameter, round and slightly flattened on one side. (4) The frame has two small holes drilled in it; one at *E* is the size of the needle, the other at *F* is somewhat larger, but has a small piece of brass with a hole the size of the needle soldered over it; this allows for a

finer adjustment. (5) The needle carrying the writing point is an ordinary darning needle 7.5 cm. long and 1 mm. in diameter. A notch is cut midway in the side of the needle to engage the trigger, and the point is blunted to give a mark of the proper size. The point is ground flat on the side which slides against the flattened edge of the dermic pencil, and the latter is then carefully ground and polished until the center



of the two points, needle and pencil, are exactly one-half millimeter apart. It is then marked on the flattened side so that it can be readily placed in its proper position. When

this is done a second dermic pencil may be fitted to the same needle, one needle being sufficient for each instrument. (6) A spring made of very thin aluminum and acting as a trigger is riveted to the frame at *D*; it has a square hole through which the needle passes, the edge of this hole engages the notch in the needle when set. (7) A stop at the end of the trigger spring prevents the back edge of the square hole from interferring with the drop of the needle. (8) A short brass tube soldered along the edge of the frame just below *C*, carries a rubber tube and 5 cm. bulb. A small hole bored through the frame at the bottom of the brass tube, opposite the centre of the needle spring, allows a jet of air from the bulb to move the trigger and release the needle.

In order to prevent the distance between the centers of the dermic pencil and marking needle from varying, the holes through which the needle slides must be drilled so that the needle fits snugly. The needle must then be moved up and down until it slides readily. No oil has been found which improves the sliding qualities; while the moisture from the fingers on the needle will interfer with it. The needle must be kept in vaseline and polished and cleaned before using.

When properly adjusted, the needle never fails to write. It makes a fine point on the skin. The mark is one-half millimeter from the localized point, always in one direction only, so that the pattern will remain the same. The eyes are not moved from the skin and the reaction time is reduced to a minimum. There are no figures at present available to demonstrate that the instrument gives a more accurate localization than the old method.

PSYCHOLOGICAL REVIEW PUBLICATIONS

Original contributions and discussions intended for the Psychological Review should be addressed to

Dr. John B. Watson, Editor PSYCHOLOGICAL REVIEW,
244 Madison Avenue, New York

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor Madison Bentley,
Editor JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
University of Illinois, Urbana, Ill.

Contributions intended for the Psychological Monographs should be addressed to

Dr. S. I. Franz, Editor PSYCHOLOGICAL MONOGRAPHS,
University of California, Sonthern Branch, Los Angeles, California

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor S. W. Fernberger, Editor PSYCHOLOGICAL BULLETIN,
University of Pennsylvania, Philadelphia

Titles and reprints intended for the Psychological Index should be sent to

Professor Walter S. Hunter, Editor PSYCHOLOGICAL INDEX,
Clark University, Worcester, Mass.

All business communications should be addressed to

Psychological Review Company,
Prince and Lemon Streets, Lancaster, Pa., and Princeton, New Jersey

---

PHILOSOPHICAL MONOGRAPHS

VOL. I

1. **Æsthetic Experience: Its Nature and Function in Epistemology.** W. D. FURRY. Pp. xvi + 160. $1.60. 2. **The Philosophy of John Norris of Bemerton.** FLORA ISABEL MACKINNON. Pp. iii + 103. $1.00. 3. **Freedom and Purpose. An Interpretation of the Psychology of Spinoza.** JAMES H. DUNHAM. Pp. iv + 126. $1.25.

LIBRARY OF GENETIC SCIENCE AND PHILOSOPHY

Vol. 1. The Moral Life. A Study in Genetic Ethics. By ARTHUR ERNEST DAVIES, PH.D., Professor of Philosophy in Colorado College. xiv + 188 pages. 8vo. Cloth. Price, $1.50 net. Postage, 10 cents.

Vol. 2. Darwin and the Humanities. By JAMES MARK BALDWIN. viii + 118 pages. 8vo. Cloth. Price, $1.50 net. Postage, 8 cents.

Published by

PSYCHOLOGICAL REVIEW COMPANY

ExP

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Cornell University.
Subscription $6.50. 624 pages annually. Edited by E. B. Titchener.
Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.; Clark Univ. Sub. $7. 700 pages ann. Ed. by Carl Murchison and an international coöperating board. Quarterly. Child behavior differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company.
Subscription $5.00. 480 pages annually.
Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company.
Subscription $5.50. 720 pages annually. Psychological literature.
Monthly. Founded 1904. Edited by Samuel W. Fernberger.

Psychological Monographs—Princeton, N. J.; Psychological Review Company.
Subscription $6.00 per vol. 500 pp. Founded 1895. Edited by Shepherd I. Franz.
Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.; Psychological Review Company.
Subscription $2.00. 200 pp. Founded 1895. Edited by Madison Bentley.
An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116thSt.
Subscription $4. 728 pages per volume. Founded 1904.
Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush and H. W. Schneider.

Archives of Psychology—Sub-Station 84, N. Y.; Archives of Psychology.
Subscription $5. 500 pp. per vol. Founded 1906. Edited by R. S. Woodworth.
Published without fixed dates, each number a single experimental study.

Journal of Abnormal Psychology and Social Psychology—Albany, N. Y.
Sub. $5. Boyd Printing Co. Ed. by Morton Prince, in cooperation with Floyd H. Allport. Quarterly. 432 pages annually. Founded 1906. Abnormal and social.

Psychological Clinic—Philadelphia; Psychological Clinic Press.
Subscription $2.50. 288 pages. Edited by Lightner Witmer. Founded 1907.
Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.

Training School Bulletin—Vineland, N. J.; The Training School.
Subscription $1. 160 pages ann. Edited by E. R. Johnstone. Founded 1904.
Monthly (10 numbers). Psychology and training of defectives.

Comparative Psychology Monographs—Baltimore; Williams and Wilkins Co.
Subscription $5. 500 pages per volume. Edited by W. S. Hunter.
Published without fixed dates, each number a single research.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W.
Subscription $6. 500 pages annually. Psychoanalysis.
Quarterly. Founded 1913. Edited by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J.; Psychological Review Company.
Subscription $5.00. 480 pages annually. Experimental.
Bi-monthly. Founded 1916. Edited by John B. Watson.

Journal of Applied Psychology—Bloomington, Ind.; Indiana University Press.
Subscription $4. 400 pages annually. Founded 1917.
Quarterly. Edited by James P. Porter and William F. Book.

Journal of Comparative Psychology—Baltimore; Williams and Wilkins Company.
Subscription $5. 500 pages annually. Founded 1921.
Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

Genetic Psychology Monographs—Worcester, Mass.; Clark University.
Subscription $7.00. 600 pp. ann. Ed. by Carl Murchison.
Bi-monthly. Each number a complete research. Child behavior, differential and genetic psychology. Founded 1925.

Ex.P.